

THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Méthodes de régularisation en programmation semi-définie

Grisard, Hélène

Award date:
2003

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

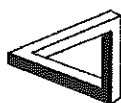
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



FUNDP
Faculté des Sciences
Département de Mathématique

Rempart de la Vierge, 8
B-5000 Namur Belgique

Méthodes de régularisation en programmation semi-définie



Mémoire présenté pour l'obtention
du grade de
Licencié en Sciences Mathématiques
Par

Grisard Hélène

Promoteur : Jean-Jacques Strodiot

Année Académique 2002-2003

Remerciements.

Je remercie mes proches pour leur soutien et leur confiance ainsi que tous ceux qui ont contribué de près ou de loin à la réalisation de ce mémoire.

Résumé.

Le but de ce mémoire est de résoudre des problèmes de programmation semi-définie en exprimant les conditions d'optimalité de deux nouvelles manières. Dans un premier temps, nous utiliserons deux fonctions de régularisation pour garder la propriété de différentiabilité : la fonction minimum et la fonction de Fischer-Burmesiter. Nous appliquerons alors à ce nouveau système une méthode de type Newton. Nous montrerons que l'algorithme engendré est bien défini, que sa convergence est globale et que sa convergence locale est superlinéaire. Nous donnerons également une formulation matrice-vecteur et une formulation de Lyapunov des systèmes de Newton.

Dans un second temps, nous considérerons les systèmes linéaires. Nous transformerons de la même manière leurs conditions d'optimalité en un système régulier grâce à des formulations régulières des fonctions minimum et de Fischer-Burmeister et nous appliquerons une méthode de type Newton. L'algorithme engendré sera bien défini et nous obtiendrons la convergence globale. Nous établirons la convergence locale quadratique pour la fonction minimum. En ce qui concerne la fonction de Fischer-Burmeister, la question de la convergence locale constitue encore une question ouverte.

Abstract.

The aim of this memory is to solve problems of semi-definite programming by formulating the optimality conditions in two new manners.

Firstly, we will use two smoothed functions to keep the differentiability property : the minimum function and the Fischer-Burmesiter function. We will then apply to this new system a Newton-type method. We will show that the generated algorithm is well defined, that its convergence is global and that its local convergence is superlinear. We will also give a matrix-vector formulation and a Lyapunov formulation of the Newton systems.

In a second time, we will consider the linear systems. We will transform their optimality conditions into a smooth system thanks to smooth formulations of the minimum functions and of the Fischer-Burmeister functions. We will then apply a Newton-type method. The generated algorithm will be well defined and we will obtain global convergence. We will establish quadratic local convergence for the minimum function. But for the Fischer-Burmeister function, the question of local convergence is still open .

Table des matières

| | | |
|-----------|---|-----------|
| I | Le cas des systèmes semi-définis positifs. | 4 |
| 1 | Introduction. | 5 |
| 2 | Fonctions régulières minimum et de Fischer-Burmeister. | 8 |
| 2.1 | Définitions et liens avec les conditions d'optimalité ou les conditions de la trajectoire centrale. | 8 |
| 2.2 | Propriétés de ϕ | 15 |
| 3 | Description de l'algorithme. | 23 |
| 4 | Convergence, Calcul du point de départ et Critère d'arrêt. | 35 |
| 4.1 | Convergence globale. | 35 |
| 4.2 | Convergence locale. | 37 |
| 4.3 | Calcul du point de départ et Critère d'arrêt. | 44 |
| 5 | Quelques aspects pratiques d'une méthode de type Newton pour les SPDs. | 46 |
| 5.1 | Formulation de Lyapunov des systèmes de Newton. | 46 |
| 5.2 | Formulation Matrice-Vecteur des systèmes de Newton. | 52 |
| 5.3 | Résultats pour la fonction de Fischer-Burmeister. | 57 |
| II | Le cas des systèmes linéaires. | 60 |
| 1 | Les méthodes régulières Prédicteur-Correcteur pour la solution des Programmes linéaires. | 61 |
| 1.1 | Introduction. | 61 |
| 1.2 | Fonctions et trajectoires régulières. | 62 |
| 1.3 | Méthode régulière Prédicteur-Correcteur. | 69 |
| 1.4 | Convergence. | 75 |

| | | |
|----------|---|------------|
| 1.4.1 | Convergence globale. | 75 |
| 1.4.2 | Convergence locale. | 79 |
| 1.5 | Calcul du point de départ et Critère d'arrêt. | 80 |
| 2 | Les méthodes régulières Prédicteur-Correcteur pour les programmes linéaires avec une mise à jour plus flexible du paramètre de régularisation. | 81 |
| 2.1 | Convergence. | 87 |
| 2.2 | Choix du point de départ et Critère d'arrêt. | 95 |
| 3 | Méthodes de type régulières améliorées pour la solution de programmes linéaires. | 96 |
| 3.1 | Description de l'algorithme. | 96 |
| 3.2 | Propriétés de convergence. | 102 |
| 3.3 | Choix du point de départ et Critère d'arrêt. | 107 |
| A | Quelques définitions et résultats. | 109 |
| B | Transformation de matrices en vecteurs. | 112 |
| C | Tableaux comparatifs pour les trois méthodes linéaires. | 114 |

Première partie

Le cas des systèmes
semi-définis positifs.

Chapitre 1

Introduction.

Dans ce travail, nous allons décrire un algorithme pour la solution des programmes semi-définis. Un programme semi-défini est un problème d'optimisation avec contraintes dont la forme *primale* est la suivante :

$$P \equiv \begin{cases} \min & C \bullet X \\ \text{s.c.} & A_i \bullet X = b_i \quad \forall i = 1, \dots, m \\ & X \succeq 0 \end{cases} \quad (1.1)$$

Le problème *dual* correspondant :

$$D \equiv \begin{cases} \max & b^T \lambda \\ \text{s.c.} & \sum_{i=1}^m \lambda_i A_i + S = C \\ & S \succeq 0 \end{cases} \quad (1.2)$$

Dans ces problèmes, les données sont le vecteur $b \in \mathbb{R}^m$ et les matrices symétriques $C \in \mathbb{R}^{n \times n}$ et $A_i \in \mathbb{R}^{n \times n} (i = 1, \dots, m)$. Par ailleurs, la matrice symétrique $X \in \mathbb{R}^{n \times n}$ désigne la variable du primal tandis que le vecteur $\lambda \in \mathbb{R}^m$ et la matrice symétrique $S \in \mathbb{R}^{n \times n}$ sont les variables du dual. On voit aisément que le problème primal est convexe.

Les conditions d'optimalité sont :

$$\begin{aligned} A_i \bullet X &= b_i \quad \forall i = 1, \dots, m \\ \sum_{i=1}^m \lambda_i A_i + S &= C \\ X \succeq 0, S \succeq 0 \text{ et } XS &= 0 \end{aligned} \quad (1.3)$$

Le problème (1.1) et son dual, le problème (1.2), ont une solution si et seulement si les conditions d'optimalité (1.3) en ont une.

Cette équivalence est une conséquence du Théorème de Dualité Forte :
L'admissibilité du primal et du dual entraîne $XS = 0$.
Généralement, la méthode utilisée pour résoudre ces conditions est la *méthode des points intérieurs*. Cette méthode considère la perturbation suivante des conditions d'optimalité :

$$\begin{aligned} \sum_{i=1}^m \lambda_i A_i + S &= C, \\ A_i \bullet X &= b_i \quad \forall i = 1, \dots, m \\ X &\succ 0, S \succ 0, XS = \tau^2 I \end{aligned} \tag{1.4}$$

où τ est un paramètre positif. Ces équations sont appelées *conditions de la trajectoire centrale*. La perturbation porte sur la condition de complémentarité, i.e., $XS = 0$.

Les méthodes des points intérieurs appliquent une approche de type Newton à ces équations.

La méthode discutée ici est une approche de type régulière. Ces méthodes consistent en une reformulation différentiable des conditions d'optimalité. Ces dernières deviennent alors un système non linéaire d'équations ne contenant pas de contraintes d'inégalité explicites comme $X \succeq 0$, $S \succeq 0$ ou $X \succ 0$. Pour arriver à un tel système, nous utilisons soit la fonction régulière de Fischer-Burmeister, soit la fonction régulière minimum afin d'obtenir un système différentiable. Ensuite, nous appliquons la méthode de Newton à ce système. Nous devons donc résoudre à chaque itération un système linéaire du type : $\nabla \Theta(W)(\Delta W) = -\Theta(W)$ où Θ est une fonction différentiable définie plus loin. Nous verrons aussi que la matrice des coefficients de ce système possède quelques propriétés remarquables. En effet, si nous utilisons la fonction régulière de Fischer-Burmeister, la matrice est définie positive. Par ailleurs, si nous prenons la fonction régulière minimum, la matrice sera toujours définie positive mais en plus symétrique. Le coût du calcul, dans le cas de la fonction de Fischer-Burmeister, sera plus élevé par le manque de symétrie.

La méthode de Newton génère automatiquement des directions de recherches symétriques sans qu'aucune transformation supplémentaire ne soit nécessaire. C'est un point de divergence avec les méthodes des points intérieurs. En effet, ces dernières doivent rendre les conditions (1.4) symétriques afin de produire des directions symétriques.

Par rapport au choix du point de départ, la méthode développée sera moins restrictive que celle des points intérieurs qui requiert un point de départ

strictement admissible. Nous présenterons de nouvelles caractérisations pour les conditions de la trajectoire centrale lesquelles peuvent être vues comme la base de notre méthode.

Dans ce texte, nous utilisons les notations suivantes :

Pour deux matrices $A, B \in \mathbb{R}^{n \times n}$, nous définissons le produit scalaire comme : $A \bullet B := \langle A, B \rangle := \text{tr}(AB^T)$ où $\text{tr}(C) := \sum_{i=1}^n c_{ii}$ est la trace d'une matrice $C \in \mathbb{R}^{n \times n}$.

Les notations $\mathcal{S}^{n \times n}$, $\mathcal{S}_+^{n \times n}$ et $\mathcal{S}_{++}^{n \times n}$ correspondent à l'ensemble des matrices symétriques, symétriques semi-définies positives et symétriques définies positives respectivement de dimension $n \times n$. Nous écrivons aussi $A \succeq 0$ et $A \succ 0$ pour indiquer que A appartient à $\mathcal{S}_+^{n \times n}$ et $\mathcal{S}_{++}^{n \times n}$ respectivement. De plus, $A \succeq B$ ou $A \succ B$ signifie que $A - B \succeq 0$ ou $A - B \succ 0$. Si $A \succeq 0$, nous désignons par $A^{1/2}$ l'unique racine carrée semi-définie positive de A .

Nous munissons l'espace vectoriel $\mathbb{R}^{n \times n} \times \mathbb{R}^m \times \mathbb{R}^{n \times n}$ de la norme

$$|||(X, \lambda, S)||| := \sqrt{\|X\|_F^2 + \|\lambda\|_2^2 + \|S\|_F^2}.$$

Nous utilisons le même symbole pour la norme

$$|||(X, \lambda, S, \tau)||| := \sqrt{\|X\|_F^2 + \|\lambda\|_2^2 + \|S\|_F^2 + \tau^2}$$

dans l'espace vectoriel $\mathbb{R}^{n \times n} \times \mathbb{R}^m \times \mathbb{R}^{n \times n} \times \mathbb{R}$. Rappelons la définition de la norme 2 (la norme euclidienne) et de la norme de Frobénius d'une matrice $A \in \mathbb{R}^{n \times n}$:

$$\|A\|_F := \sqrt{\sum_{i=1}^n |a_{ij}|^2}$$

$$\|A\|_2 := \max_i \sqrt{\mu_i} \text{ où } \mu_i \text{ valeur propre de } A^*A, \forall i = 1, \dots, n.$$

Chapitre 2

Fonctions régulières minimum et de Fischer-Burmeister.

2.1 Définitions et liens avec les conditions d'optimalité ou les conditions de la trajectoire centrale.

Le but de cette section est d'exprimer les conditions de la trajectoire centrale (1.4) pour les SDPs de deux nouvelles façons. Ce procédé sera possible à l'aide des fonctions minimum et de Fischer-Burmeister. Il me semble important de souligner que nous utiliserons une formulation régulière de ces fonctions pour que la propriété de différentiabilité soit satisfaite.

Pour motiver notre approche, définissons la fonction $\varphi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ par

$$\varphi(a, b) := a + b - \sqrt{a^2 + b^2}.$$

Cette fonction fut introduite par Fischer et est habituellement appelée la fonction de Fischer-Burmeister. Cette fonction possède la propriété suivante:

$$\varphi(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0 \text{ et } ab = 0.$$

Définissons ensuite la fonction $\phi : \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$ par

$$\phi(X, S) := X + S - (X^2 + S^2)^{1/2}. \quad (2.1)$$

Cette fonction est une extension évidente de la définition de φ ; les arguments sont deux matrices symétriques plutôt que deux nombres réels. Montrons

ensuite que la fonction ϕ vérifie une propriété similaire à celle de φ , i.e.,

$$\phi(X, S) = 0 \Leftrightarrow X \succeq 0, S \succeq 0 \text{ et } XS = 0. \quad (2.2)$$

Pour rendre la preuve plus aisée, nous démontrons d'abord le lemme suivant :

Lemme 2.1 Soient $X, S \in S_+^{n \times n}$.
Alors $XS = 0$ si et seulement si $X \bullet S = 0$.

Preuve :

Supposons d'abord que $XS = 0$. Alors, $X \bullet S = \text{tr}[XS] = 0$.

Supposons ensuite que $X \bullet S = 0$. Comme $S \succeq 0$, il suit du théorème de décomposition spectrale que $S = U\Omega U^T$ où U est une matrice orthogonale et $\Omega = \text{diag}(\omega_i)$ et $\omega_i \geq 0$ pour tout $i = 1, \dots, n$. Alors, $C := U^T X U$ est symétrique semi-définie positive car elle est semblable à la matrice X . En particulier, ses éléments diagonaux sont tels que $c_{ii} \geq 0$. Nous devons seulement montrer que $C\Omega = 0$. En effet, de $X \bullet S = 0$, nous avons $C \bullet \Omega = 0$ et, par conséquent, $\sum_{i=1}^n c_{ii}\omega_i = 0$. Puisque tous les éléments de cette somme sont non négatifs, il suit qu'ils sont tous égaux à zéro. Montrons ensuite que $c_{ij}w_j = 0$ pour $i \neq j$. Si ce n'est pas le cas, $c_{ij}w_j \neq 0$ et $w_j > 0$. Par conséquent, $c_{jj} = 0$ et, la matrice C étant symétrique semi-définie positive, la ligne j et la colonne j sont nulles. (cfr [9, page 147]). Mais, alors, $c_{ij} = 0$ ce qui contredit $c_{ij}w_j \neq 0$. La matrice $C\Omega$ est donc nulle. La thèse en découle. \square

Proposition 2.1 Soit ϕ la fonction de Fischer-Burmeister définie en (2.1).
Alors

$$\phi(X, S) = 0 \Leftrightarrow X \succeq 0, S \succeq 0, XS = 0.$$

Preuve :

Supposons pour commencer que $X \succeq 0, S \succeq 0$ et $XS = 0$. Alors $XS + SX = 0$ et $(X + S)^2 = X^2 + S^2$. En utilisant le fait que X et S sont symétriques semi-définies positives, nous obtenons

$$X + S = (X^2 + S^2)^{1/2},$$

car la racine carrée d'une matrice symétrique semi-définie positive est définie uniquement dans l'espace des matrices symétriques semi-définies positives.

Nous en déduisons $\phi(X, S) = 0$.

Supposons ensuite que $\phi(X, S) = 0$ soit satisfaite pour deux matrices $X, S \in \mathcal{S}^{n \times n}$, i.e.,

$$X + S = (X^2 + S^2)^{1/2}.$$

En élevant au carré les deux membres de cette expression, nous avons

$$X^2 + S^2 = (X + S)^2 \text{ et } X + S \in \mathcal{S}_+^{n \times n}.$$

Ce qui est équivalent à

$$XS + SX = 0 \text{ et } X + S \in \mathcal{S}_+^{n \times n}. \quad (2.3)$$

Soit $X = Q^T D Q$ avec $Q \in \mathbb{R}^{n \times n}$ matrice orthogonale et $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ la décomposition spectrale de la matrice symétrique X . Les relations (2.3) peuvent alors être mises sous la forme :

$$Q^T D Q S + S Q^T D Q = 0 \text{ et } Q^T D Q + S \in \mathcal{S}_+^{n \times n}.$$

Prémultipliant par Q et postmultipliant par Q^T , nous avons

$$D Q S Q^T + Q S Q^T D = 0 \text{ et } D + Q S Q^T \in \mathcal{S}_+^{n \times n}.$$

Posant $A := Q S Q^T$, nous obtenons :

$$D A + A D = 0 \text{ et } D + A \in \mathcal{S}_+^{n \times n}.$$

Du point de vue des composantes, nous pouvons écrire :

$$(\lambda_i + \lambda_j) a_{ij} = 0 \text{ et } D + A \in \mathcal{S}_+^{n \times n} \quad \forall i, j = 1, \dots, n.$$

Si nous prenons $i = j$, nous avons $2\lambda_i a_{ii} = 0$ et $\lambda_i + a_{ii} \geq 0$ pour tout $i = 1, \dots, n$. Il suit que $\lambda_i \geq 0$ pour tout $i = 1, \dots, n$, i.e., X est semi-définie positive. Par un argument analogue, nous arrivons à la conclusion que S est semi-définie positive. Il reste à montrer que $XS = 0$. Pour cela, remarquons que

$$(2.3) \Rightarrow X \bullet S = \text{tr}[XS] = \frac{1}{2} \text{tr}[XS + SX] = 0.$$

Par le lemme 2.1, nous concluons que $XS = 0$.

□

Nous voulons modifier la définition de la fonction ϕ afin de l'utiliser pour caractériser les conditions de la trajectoire centrale. Pour cela, considérons τ un nombre quelconque non négatif vu dans un premier temps comme un paramètre. Définissons $\varphi_\tau : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ par

$$\varphi_\tau(a, b) := a + b - \sqrt{a^2 + b^2 + 2\tau^2}.$$

Cette fonction est appelée fonction régulière de Fischer-Burmeister (elle est continûment différentiable pour chaque $\tau > 0$). Nous constatons que lorsque $\tau = 0$, elle coïncide avec la fonction de Fischer-Burmeister. La fonction φ_τ possède la propriété suivante :

$$\varphi_\tau(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0, ab = \tau^2.$$

Maintenant, nous généralisons la fonction régulière de Fischer-Burmeister φ_τ de la façon suivante : définissons $\phi_\tau : \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$ par

$$\phi_\tau(X, S) := X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2}. \quad (2.4)$$

Nous établissons le résultat suivant :

Proposition 2.2 Soient $\tau > 0$ et $\phi_\tau(X, S) := X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2}$. Alors

$$\phi_\tau(X, S) = 0 \Leftrightarrow X \succ 0, S \succ 0 \text{ et } XS = \tau^2 I.$$

Preuve :

Supposons d'abord que $X \succ 0, S \succ 0$ et $XS = \tau^2 I$. Alors, $XS + SX = 2\tau^2 I$ car $XS = (XS)^T = S^T X^T = SX$ et par conséquent,

$$(X + S)^2 = X^2 + S^2 + 2\tau^2 I.$$

Du fait de la symétrie et de la définie positivité de X et de S , il suit que

$$X + S = (X^2 + S^2 + 2\tau^2 I)^{1/2}.$$

Nous obtenons ainsi la thèse, i.e., $\phi_\tau(X, S) = 0$.

Supposons ensuite que $\phi_\tau(X, S) = 0$ pour deux matrices $X, S \in \mathcal{S}^{n \times n}$.

Alors, $X + S = (X^2 + S^2 + 2\tau^2 I)^{1/2}$. En prenant le carré des deux côtés de l'équation, nous avons

$$X^2 + S^2 + 2\tau^2 I = (X + S)^2 \text{ et } X + S \in \mathcal{S}_{++}^{n \times n},$$

ce qui est équivalent à

$$XS + SX = 2\tau^2 I \text{ et } X + S \in \mathcal{S}_{++}^{n \times n}. \quad (2.5)$$

Soit $X = Q^T D Q$ avec $Q \in \mathbb{R}^{n \times n}$ orthogonale et $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, la décomposition spectrale de la matrice symétrique X . Par la preuve de la proposition précédente et en utilisant $A := Q S Q^T$, (2.5) peut s'écrire comme

$$DA + AD = 2\tau^2 I \text{ et } D + A \in \mathcal{S}_{++}^{n \times n}. \quad (2.6)$$

Du point de vue des composantes, nous obtenons

$$(\lambda_i + \lambda_j)a_{ij} = 2\tau^2 \delta_{ij} \text{ et } D + A \in \mathcal{S}_{++}^{n \times n} \quad \forall i, j \in \underline{n}. \quad (2.7)$$

Si $i = j$, nous avons $2\lambda_i a_{ii} = 2\tau^2$ et $\lambda_i + a_{ii} > 0$ pour tout $i \in \underline{n}$. D'où, $\lambda_i > 0$ pour tout $i \in \underline{n}$ et donc, X est une matrice symétrique définie positive.

Par un argument similaire, nous prouvons que S est une matrice symétrique définie positive.

Pour vérifier que $XS = \tau^2 I$, nous constatons que (2.7) implique que $a_{ij} = 0$ pour tout $i \neq j$ puisque $\lambda_i + \lambda_j > 0$. La matrice A est par conséquent une matrice diagonale. En particulier, nous avons $DA = AD$ car deux matrices diagonales commutent. De (2.6), nous obtenons $DA = \tau^2 I$. En prémultipliant par Q^T et en postmultipliant par Q , nous déduisons $XS = Q^T D Q S = Q^T D A Q = \tau^2 I$.

□

Nous voulons à présent introduire une seconde fonction avec des propriétés semblables à celles de la fonction (régulière) de Fischer-Burmeister. A cette fin, soit

$$\varphi(a, b) := 2 \min\{a, b\}$$

pour $a, b \in \mathbb{R}$. Pour des raisons évidentes, cette fonction est appelée fonction minimum. Il est aisé de constater qu'elle vérifie l'équivalence

$$\varphi(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0 \text{ et } ab = 0.$$

Afin d'étendre cette définition à la classe des matrices symétriques, il est judicieux de formuler la fonction minimum de cette façon :

$$\varphi(a, b) = 2 \min\{a, b\} = a + b - |a - b| = a + b - \sqrt{(a - b)^2}.$$

Grâce à cette dernière formulation, nous sommes en mesure de définir la fonction $\phi: \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$ par

$$\phi(X, S) := X + S - ((X - S)^2)^{1/2}. \quad (2.8)$$

Il s'avère que cette fonction possède également la propriété reprise dans la proposition ci-dessous :

Proposition 2.3 *Soit ϕ la fonction minimum définie en (2.8). Alors*

$$\phi(X, S) = 0 \Leftrightarrow X \succeq 0, S \succeq 0, XS = 0.$$

Preuve :

Supposons d'abord que $X \succeq 0, S \succeq 0$ et $XS = 0$. Par le lemme 2.1, $XS = 0$ est équivalent à $X \bullet S = 0$. Dès lors, pour tout $A \in \mathcal{S}_+^{n \times n}$, nous avons

$$\|(X - S) - A\|^2 = \|X - A\|^2 + 2((A - X) \bullet S) + \|S\|^2 = \|X - A\|^2 + 2(A \bullet S) + \|S\|^2.$$

Le membre de droite atteint son minimum en $A = X$. D'où,

$$X = \operatorname{argmin}_{A \in \mathcal{S}_+^{n \times n}} \|(X - S) - A\| = X + (1/2)(|X - S| - (X + S))$$

et donc, $\phi(X, S) = |X - S| - (X + S) = 0$.

Supposons ensuite que $\phi(X, S) = 0$, i.e., $X = \operatorname{argmin}_{A \in \mathcal{S}_+^{n \times n}} \|(X - S) - A\|$.

Alors $X \in \mathcal{S}_+^{n \times n}$ et

$$0 \leq \|(X - S) - A\|^2 - \|S\|^2 = \|X - A\|^2 + 2(A - X) \bullet S \quad \forall A \in \mathcal{S}_+^{n \times n}. \quad (2.9)$$

Par conséquent, pour tout $D \in \mathcal{S}_+^{n \times n}$ et pour tout $t > 0$, nous avons $A := X + tD \in \mathcal{S}_+^{n \times n}$ et (2.9) devient

$$0 \leq t^2 \|D\|^2 + 2t(D \bullet S).$$

Si nous divisons les deux membres par t et si nous faisons tendre $t \rightarrow 0$ alors

$$0 \leq D \bullet S \quad \forall D \in \mathcal{S}_+^{n \times n}.$$

Par la relation (11) de la référence [17], nous obtenons $S \in \mathcal{S}_+^{n \times n}$.

De manière similaire, pour tout $t \in (0, 1]$ et en posant $A := (1 - t)X \in \mathcal{S}_+^{n \times n}$, (2.9) devient

$$0 \leq t^2 \|X\|^2 - 2t(X \bullet S).$$

Divisant alors les deux membres par t et faisant tendre $t \rightarrow 0$, nous obtenons

$$0 \leq -(X \bullet S).$$

D'autre part, $X, S \in \mathcal{S}_+^{n \times n}$, d'où, par l'équation (11) de la référence [17], $X \bullet S \geq 0$. D'où $X \bullet S = 0$. □

Nous modifions la définition de la fonction minimum de façon à obtenir une caractérisation de la trajectoire centrale (1.4). Pour y parvenir, nous considérons la fonction minimum suivante

$$\varphi_\tau(a, b) := a + b - \sqrt{(a - b)^2 + 4\tau^2},$$

où τ désigne un nombre non négatif. Pour chaque $\tau > 0$, cette fonction possède la propriété :

$$\varphi_\tau(a, b) = 0 \Leftrightarrow a > 0, b > 0 \text{ et } ab = \tau^2.$$

Cette constatation nous pousse à définir une fonction $\phi_\tau : \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n}$ par

$$\phi_\tau(X, S) := X + S - ((X - S)^2 + 4\tau^2 I)^{1/2}. \quad (2.10)$$

Cette fonction possède la propriété :

Proposition 2.4 *Soient $\tau > 0$ et ϕ définie par (2.10). Alors*

$$\phi_\tau(X, S) = 0 \Leftrightarrow X \succ 0, S \succ 0 \text{ et } XS = \tau^2 I.$$

Preuve :

Elle est analogue à celle de la Proposition 2.2. □

Dans la suite, ϕ_τ désignera soit la fonction régulière de Fischer-Burmeister (2.4) soit la fonction régulière minimum (2.10). Nous définissons ainsi la fonction $\Phi_\tau : \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \rightarrow \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$ par

$$\Phi_\tau(X, \lambda, S) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi_\tau(X, S) \end{pmatrix}. \quad (2.11)$$

Les propositions 2.2 et 2.4 donnent immédiatement la nouvelle caractérisation des conditions de la trajectoire centrale (1.4) pour les SDPs.

Théorème 2.1 *Soit Φ_τ définie en (2.11), avec ϕ donné par (2.4) ou (2.10) et soit $\tau > 0$. Alors les énoncés suivants sont équivalents :*

(a) (X, λ, S) vérifie les conditions de la trajectoire centrale.

(b) $\Phi_\tau(X, \lambda, S) = 0$.

2.2 Propriétés de ϕ .

Dans cette section, nous montrerons que les fonctions ϕ_τ sont différentiables (au sens de Fréchet). Nous considérons à partir de maintenant le nombre non négatif τ comme une variable indépendante. Pour une question de clarté, nous noterons $\phi(X, S, \tau) := \phi_\tau(X, S)$, i.e., l'expression suivante désignera la fonction régulière de Fischer-Burmeister

$$\phi(X, S, \tau) := X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2} \quad (2.12)$$

et celle-ci, la fonction régulière minimum

$$\phi(X, S, \tau) := X + S - ((X - S)^2 + 4\tau^2 I)^{1/2}. \quad (2.13)$$

Pour notre analyse des propriétés des fonctions ϕ , nous avons besoin des deux résultats suivants :

Lemme 2.2 *Soit ϕ une des deux fonctions définies en (2.12) et (2.13). Alors, pour tous $X, S \in \mathcal{S}^{n \times n}$ et pour tous $\tau > \nu > 0$, nous avons :*

(a) $\kappa(\tau - \nu)I \succeq \phi(X, S, \nu) - \phi(X, S, \tau) \succ 0$,

(b) $\kappa\tau I \succeq \phi(X, S, 0) - \phi(X, S, \tau) \succ 0$,

où κ est une constante positive indépendante de X, S, τ et ν .

Une petite parenthèse pour préciser que la constante κ est connue. En effet, $\kappa = 2$ pour la fonction régulière minimum et $\kappa = \sqrt{2}$ pour la fonction régulière de Fischer-Burmeister.

Preuve :

Supposons que ϕ soit donné par (2.12). Soit $A = X^2 + S^2$. Considérons ensuite Q une matrice orthogonale et $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ satisfaisant à

$$A = Q^T \text{diag}[\lambda_1, \dots, \lambda_n] Q.$$

Alors,

$$\begin{aligned} \phi(X, S, \tau) &= X + S - (A + 2\tau^2 I)^{1/2} \\ &= X + S - Q^T \text{diag}[(\lambda_1 + 2\tau^2)^{1/2}, \dots, (\lambda_n + 2\tau^2)^{1/2}] Q. \end{aligned}$$

Nous écrivons $\phi(X, S, \nu)$ de façon semblable. D'où

$$\phi(X, S, \nu) - \phi(X, S, \tau) = Q^T \text{diag}[(\lambda_i + 2\tau^2)^{1/2} - (\lambda_i + 2\nu^2)^{1/2}]_{i=1}^n Q.$$

Ainsi, comme $A \succeq 0$ et $\lambda_i \geq 0$ pour tout i , nous avons

$$0 < (\lambda_i + 2\tau^2)^{1/2} - (\lambda_i + 2\nu^2)^{1/2} \leq \sqrt{2}(\tau - \nu)$$

où la seconde inégalité vient du fait que pour tout $\lambda \in \mathbb{R}_+$, la fonction $h(\tau) := (\lambda + 2\tau^2)^{1/2}$ est convexe différentiable sur \mathbb{R}_{++} avec $h'(\tau) = 2(\lambda + 2\tau^2)^{-1/2}\tau \leq \sqrt{2}$.

Nous avons alors $h(\tau) - h(\nu) \leq h'(\tau)(\tau - \nu) \leq \sqrt{2}(\tau - \nu)$ et donc

$$\sqrt{2}(\tau - \nu)I \succeq \phi(X, S, \nu) - \phi(X, S, \tau) \succ 0.$$

La première relation est satisfaite avec $\kappa = \sqrt{2}$.

Puisque la relation (a) est satisfaite pour tout $\nu \in (0, \tau)$, faisons tendre ν vers 0.

$$\begin{aligned} \phi(X, S, 0) &= \lim_{\tau \rightarrow 0} \phi(X, S, \tau) \\ &= X + S - P^T \text{diag}[\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}] P \\ &= X + S - (X^2 + S^2)^{1/2}. \end{aligned}$$

La preuve pour la fonction (2.13) est similaire.

□

Corollaire 2.1 Soient ϕ donnée par (2.12) ou (2.13) et κ la constante du lemme 2.2. Alors, les inégalités suivantes sont vérifiées :

- (a) $\|\phi(X, S, \nu) - \phi(X, S, \tau)\|_F \leq \kappa\sqrt{n}(\tau - \nu)$ pour tous $X, S \in \mathcal{S}^{n \times n}$ et pour tous $\tau > \nu > 0$.
- (b) $\|\phi(X, S, 0) - \phi(X, S, \tau)\|_F \leq \kappa\sqrt{n}\tau$ pour tout $X, S \in \mathcal{S}^{n \times n}$ et pour tout $\tau > 0$.

Preuve :

Soient $\lambda_1, \dots, \lambda_n$ les valeurs propres de $\phi(X, S, \nu) - \phi(X, S, \tau)$. Par la relation $\kappa(\tau - \nu)I \geq \phi(X, S, \nu) - \phi(X, S, \tau)$ du lemme 2.2, nous déduisons l'équation suivante : $\kappa(\tau - \nu) \geq \lambda_i > 0$. Ainsi,

$$\|\phi(X, S, \nu) - \phi(X, S, \tau)\| = \sqrt{\lambda_1^2 + \dots + \lambda_n^2} \leq \sqrt{n}\kappa(\tau - \nu).$$

par invariance de la norme sous transformation orthogonale. Nous obtenons (a). Un raisonnement analogue est effectué pour arriver à (b). □

Nous voulons montrer que les deux fonctions ϕ définies en (2.12) et en (2.13) sont continûment différentiables en leurs arguments X, S et τ au moins avec des hypothèses adéquates. Pour cela, nous avons besoin du lemme suivant :

Lemme 2.3 Soient $A \in \mathcal{S}_{++}^{n \times n}$ et $B \in \mathcal{S}_+^{n \times n}$ deux matrices données. Alors,

$$\|A^{1/2} - B^{1/2}\|_2 \leq \|A^{-1/2}\|_2 \|A - B\|_2. \quad (2.14)$$

Preuve :

Soient $A, B \in \mathcal{S}^{n \times n}$. Alors, par [10, section 7.2],

$$\|A - B\|_2 \leq \frac{\|A^2 - B^2\|_2}{\lambda_{\min}(A) + \lambda_{\min}(B)}.$$

Appliquant cette formule aux matrices $A^{1/2}$ et $B^{1/2}$, nous avons

$$\|A^{1/2} - B^{1/2}\|_2 \leq \frac{\|A - B\|_2}{\lambda_{\min}(A^{1/2}) + \lambda_{\min}(B^{1/2})}.$$

La matrice B étant semi-définie positive, nous en déduisons

$$\|A^{1/2} - B^{1/2}\|_2 \leq \frac{\|A - B\|_2}{\lambda_{\min}(A^{1/2})}.$$

Or,

$$\frac{1}{\lambda_{\min}(A^{1/2})} = \lambda_{\max}(A^{-1/2}) = \|A^{-1/2}\|_2.$$

D'où la thèse. □

Notre but est d'arriver à une formule pour les dérivées de la fonction ϕ . Supposons que ϕ désigne la fonction régulière de Fischer-Burmeister (2.12). Nous allons montrer que l'égalité

$$\begin{aligned} & \|\phi(X + U, S + V, \tau + \mu) - \phi(X, S, \tau) - \nabla\phi(X, S, \tau)(U, V, \mu)\|_2 \\ &= o(\|(U, V, \mu)\|) \end{aligned}$$

est satisfaite pour tout $(U, V, \mu) \in \mathcal{S}^{n \times n} \times \mathcal{S}^{n \times n} \times \mathbb{R}$ tendant vers $(0, 0, 0)$, où $\nabla\phi(X, S, \tau)$ est l'opérateur linéaire adéquat désignant le gradient de ϕ au point (X, S, τ) . Afin d'y arriver, nous décomposons la fonction ϕ en $\phi(X, S, \tau) = \phi_1(X, S, \tau) - \phi_2(X, S, \tau)$ avec

$$\phi_1(X, S, \tau) := X + S, \quad \phi_2(X, S, \tau) := (X^2 + S^2 + 2\tau^2 I)^{1/2}.$$

Nous pouvons aisément constater que ϕ_1 est différentiable grâce à l'égalité

$$\nabla\phi_1(X, S, \tau)(U, V, \mu) = U + V.$$

En effet, nous avons

$$\begin{aligned} & \lim_{t \rightarrow 0} \frac{\phi_1((X, S, \tau) + t(U, V, \mu)) - \phi_1(X, S, \tau)}{t} \\ &= \lim_{t \rightarrow 0} \frac{X + tU + S + tV - X - S}{t} \\ &= U + V. \end{aligned}$$

Le cas de ϕ_2 est plus compliqué. Définissons

$$E := (X^2 + S^2 + 2\tau^2 I)^{1/2}$$

et supposons E définie positive. Soit

$$L_E[X] := EX + XE$$

l'opérateur de Lyapunov correspondant. La définie positivité de E assure que l'équation de Lyapunov $L_E[X] = H$ a une solution unique dans l'ensemble des matrices symétriques pour chaque $H \in \mathcal{S}^{n \times n}$. (cfr [11, Théorème 2.2.3]). Ainsi, l'inverse L_E^{-1} de L_E est bien défini et $L_E^{-1}[H]$ correspond à l'unique élément X vérifiant $EX + XE = H$.

De plus, nous définissons la matrice D de la sorte

$$D := ((X + U)^2 + (S + V)^2 + 2(\tau + \mu)^2 I)^{1/2}.$$

Un calcul rapide montre que $D^2 - E^2 = L_E[D - E] + (D - E)^2$.

En effet,

$$\begin{aligned} L_E[D - E] + (D - E)^2 &= E(D - E) + (D - E)E \\ &\quad + (D^2 + E^2 - ED - DE) \\ &= ED - E^2 + DE - E^2 + D^2 + E^2 \\ &\quad - ED - DE \\ &= D^2 - E^2. \end{aligned}$$

L'application de L_E^{-1} à cette équation et la réorganisation des termes donnent

$$\begin{aligned} E - D &= L_E^{-1}[(D - E)^2 - (D^2 - E^2)] \\ &= L_E^{-1}[(E - D)^2] \\ &\quad - L_E^{-1}[(XU + UX + SV + VS + 4\tau\mu I + U^2 + V^2 + 2\mu^2 I)]. \end{aligned}$$

Par la linéarité de L_E^{-1} , nous avons

$$\begin{aligned} &\phi_2(X + U, S + V, \tau + \mu) - \phi_2(X, S, \tau) - \nabla\phi_2(X, S, \tau)(U, V, \mu) \\ &= -\nabla\phi_2(X, S, \tau)(U, V, \mu) - (E - D) \\ &= -\nabla\phi_2(X, S, \tau)(U, V, \mu) + L_E^{-1}[XU + UX + SV + VS + 4\tau\mu I] \\ &\quad + L_E^{-1}[U^2 + V^2 + 2\mu^2 I] - L_E^{-1}[(E - D)^2]. \end{aligned} \tag{2.15}$$

De façon évidente, nous avons

$$\|L_E^{-1}[U^2 + V^2 + 2\mu^2 I]\|_F = O(\|(U, V, \mu)\|^2).$$

En effet,

$$\begin{aligned} \|L_E^{-1}[U^2 + V^2 + 2\mu^2 I]\|_F &\leq c\|U^2 + V^2 + 2\mu^2 I\|_F \\ &\leq c(\|U^2\|_F + \|V^2\|_F + 2\|\mu^2 I\|_F) \\ &\leq c(\|U\|_F^2 + \|V\|_F^2 + \mu^2). \end{aligned}$$

En vertu du lemme 2.3 et des propriétés d'équivalence des normes,

$$\begin{aligned} \|(E - D)^2\|_F &\leq \|E - D\|_F^2 \leq \gamma_1 \|E^2 - D^2\|_F^2 \\ &= \gamma_1 \|XU + UX + SV + VS + 4\tau\mu I + U^2 + V^2 + 2\mu^2 I\|_F^2 \\ &= O(\|(U, V, \mu)\|^2). \end{aligned}$$

pour une constante $\gamma_1 > 0$ indépendante de U, V, μ . Cela implique que

$$\|L_E^{-1}[(E - D)^2]\|_F = O(\|(U, V, \mu)\|^2).$$

Par conséquent, en posant

$$\nabla\phi_2(X, S, \tau)(U, V, \mu) := L_E^{-1}[XU + UX + SV + VS + 4\tau\mu I],$$

il suit immédiatement de (2.15) que ϕ_2 est différentiable en (X, S, τ) . Nous avons ϕ différentiable en ce point. Cela prouve la partie principale du premier énoncé du résultat suivant :

Théorème 2.2 Soient $X, S \in \mathcal{S}^{n \times n}$ deux matrices données et $\tau \in \mathbb{R}_+$.

1. Si ϕ est donné par (2.12) et $X^2 + S^2 + 2\tau^2 I \succ 0$, alors ϕ est continûment différentiable en (X, S, τ) avec :

$$\begin{aligned} \nabla \phi(X, S, \tau)(U, V, \mu) &= U + V - L_E^{-1}[XU + UX + SV + VS] \\ &\quad - L_E^{-1}[4\tau\mu I], \end{aligned} \tag{2.16}$$

où $E := (X^2 + S^2 + 2\tau^2 I)^{1/2}$.

2. Si ϕ est donné par (2.13) et $(X - S)^2 + 4\tau^2 I \succ 0$, alors ϕ est continûment différentiable en (X, S, τ) avec :

$$\begin{aligned} \nabla \phi(X, S, \tau)(U, V, \mu) &= U + V - L_E^{-1}[(X - S)(U - V)] \\ &\quad - L_E^{-1}[(U - V)(X - S) + 8\tau\mu I], \end{aligned} \tag{2.17}$$

où $E := ((X - S)^2 + 4\tau^2 I)^{1/2}$.

Preuve :

1. La différentiabilité de la fonction régulière de Fischer-Burmeister suit de la discussion précédente. Nous voyons que $\nabla \phi(X, S, \tau)$ est continu en (X, S, τ) par linéarité de l'opérateur. Nous concluons que ϕ est continûment différentiable en (X, S, τ) .

2. Nous savons que $E := ((X - S)^2 + 4\tau^2 I)^{1/2}$.

Soit

$$D := ((X - S + U - V)^2 + 4(\tau + \mu)^2 I)^{1/2}.$$

Utilisons le résultat suivant : $D^2 - E^2 = L_E[D - E] + (D - E)^2$ et appliquons L_E^{-1} à cette équation. Nous obtenons

$$\begin{aligned} E - D &= L_E^{-1}[(E - D)^2] \\ &\quad - L_E^{-1}[(X - S)(U - V) + (U - V)(X - S) + (U - V)^2] \\ &\quad + L_E^{-1}[8\tau\mu I + 4\mu^2 I]. \end{aligned}$$

ϕ est divisé en deux comme pour la première partie. La partie correspondant à ϕ_1 reste identique. Puisque L_E^{-1} est linéaire, nous avons

$$\begin{aligned}
& \phi_2(X + U, S + V, \tau + \mu) - \phi_2(X, S, \tau) - \nabla \phi_2(X, S, \tau)(U, V, \mu) \\
&= -(E - D) - \nabla \phi_2(X, S, \tau)(U, V, \mu) \\
&= -\nabla \phi_2(X, S, \tau)(U, V, \mu) \\
&\quad + L_E^{-1}[(X - S)(U - V) + (U - V)(X - S) + 8\tau\mu I] \\
&\quad - L_E^{-1}[(E - D)^2] + L_E^{-1}[(U - V)^2 + 4\mu^2 I].
\end{aligned}$$

Grâce au lemme 2.3,

$$\begin{aligned}
& \|(E - D)^2\|_F \\
&\leq \|E - D\|_F^2 \leq \|E^2 - D^2\|_F^2 \\
&\leq \gamma_2 \|(X - S)(U - V) + (U - V)(X - S) + (U - V)^2 + 8\tau\mu I + 4\mu^2 I\|_F^2 \\
&= O(\|(U, V, \mu)\|^2)
\end{aligned}$$

pour une constante $\gamma_2 > 0$ indépendante de U, V, μ .

De plus,

$$\|L_E^{-1}[(U - V)^2 + 4\mu^2 I]\|_F = O(\|(U, V, \mu)\|^2).$$

D'où, en posant

$$\nabla \phi_2(X, S, \tau)(U, V, \mu) := -L_E^{-1}[(X - S)(U - V) + (U - V)(X - S) + 8\tau\mu I]$$

nous arrivons à ϕ_2 différentiable en (X, S, τ) . Nous venons donc de montrer que ϕ est différentiable.

Puisque l'opérateur est linéaire, nous remarquons que $\nabla \phi(X, S, \tau)$ est continue en (X, S, τ) .

□

Notons que ce théorème implique que si $\tau > 0$, alors les deux fonctions ϕ sont continûment différentiables partout.

Chapitre 3

Description de l'algorithme.

Nous voulons maintenant exploiter nos résultats précédents pour obtenir un algorithme approprié pour la solution des conditions d'optimalité (1.3), et, par conséquent, pour la solution des programmes semi-définis primal dual cités ci-dessus. La façon la plus adéquate est d'utiliser la fonction suivante :

$$\Phi(X, \lambda, S) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi(X, S) \end{pmatrix}. \quad (3.1)$$

avec ϕ étant la fonction de Fischer-Burmeister (2.1) ou la fonction minimum (2.8). Les propositions 2.1 et 2.3 impliquent directement que

(X^*, λ^*, S^*) résout (1.3) ssi (X^*, λ^*, S^*) est solution de $\Phi(X, \lambda, S) = 0$.

Cependant, résoudre le système non linéaire d'équations $\Phi(X, \lambda, S) = 0$ est une tâche non triviale parce que ϕ , et par conséquent Φ , sont non différentiables en général.

L'idée est de remplacer la fonction non régulière Φ par la fonction différentiable

$$\Phi_\tau(X, \lambda, S) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi_\tau(X, S) \end{pmatrix}. \quad (3.2)$$

où ϕ_τ désigne soit la fonction régulière de Fischer-Burmeister (2.4) ou la fonction régulière minimum (2.10). Nous avons une équivalence entre le système non linéaire d'équations $\Phi_\tau(X, \lambda, S) = 0$ et les conditions de la trajectoire

centrale (1.4).

Considérons d'abord τ comme une variable indépendante. Définissons la fonction $\Theta: \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R} \rightarrow \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}$ par

$$\Theta(X, \lambda, S, \tau) := \begin{pmatrix} \sum_{i=1}^m \lambda_i A_i + S - C \\ A_i \bullet X - b_i \quad (i = 1, \dots, m) \\ \phi(X, S, \tau) \\ \tau \end{pmatrix}, \quad (3.3)$$

où ϕ représente une des deux fonctions données par (2.12) ou (2.13). Mis à part le fait que τ soit une variable indépendante plutôt qu'un paramètre, la fonction Θ diffère aussi de la fonction Φ_τ par l'ajout d'une ligne tel que

$$\Theta(X, \lambda, S, \tau) = 0 \quad (3.4)$$

est un système carré d'équations. La dernière équation, $\tau = 0$, implique directement que le système (3.4) est équivalent aux conditions d'optimalité (1.3) elles-mêmes. De plus, il suit du théorème 2.2 que Θ est une fonction continûment différentiable en chaque point (X, λ, S, τ) . La véracité de cette proposition vient du fait que τ soit strictement positif. Cette condition sera garantie par notre méthode.

L'idée principale de notre algorithme est de résoudre le système d'équations (3.4) par la *méthode de Newton*. La convergence globale de cette méthode sera atteinte en suivant un voisinage adéquat de la trajectoire centrale. Le voisinage utilisé dans ce cas est

$$\mathcal{N}(\beta) = \left\{ (X, \lambda, S, \tau) \mid \begin{aligned} &A_i \bullet X = b_i \quad \forall i = 1, \dots, m, \\ &\sum_{i=1}^m \lambda_i A_i + S = C, \|\phi(X, S, \tau)\|_F \leq \beta\tau \end{aligned} \right\}, \quad (3.5)$$

où β est un nombre positif.

Une convergence locale rapide sera assurée par un pas Prédicteur judicieusement choisi.

Afin d'alléger la notation, introduisons les abréviations suivantes :

$W^k := (X^k, \lambda^k, S^k)$ où k est l'indice d'itération. Enonçons notre méthode de type régulière pour la résolution des programmes semi-définis.

Algorithme.

Etape 0: Initialisation.

Choisir $\overline{W^0} = (X^0, \lambda^0, S^0) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n}$ avec $\sum_{i=1}^m \lambda_i^0 A_i + S^0 = C$
et $A_i \bullet X^0 = b_i$ ($i = 1, \dots, m$).

Choisir $\tau_0 > 0, \beta > 0$ avec $\|\phi(X^0, S^0, \tau_0)\|_F \leq \beta\tau_0$, et poser $k := 0$.

Choisir $\hat{\sigma}, \alpha_1, \alpha_2 \in (0, 1)$.

Etape 1: Pas Prédicteur.

Soit $(\Delta W^k, \Delta \tau_k) = (\Delta X^k, \Delta \lambda^k, \Delta S^k, \Delta \tau_k) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}$
une solution du système

$$\nabla \Theta(W^k, \tau_k) \begin{pmatrix} \Delta W \\ \Delta \tau \end{pmatrix} = -\Theta(W^k, \tau_k). \quad (3.6)$$

Si $\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F = 0$, STOP.

Autrement, si

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k)\|_F > \beta\tau_k,$$

alors poser

$$\widehat{W}^k := W^k, \widehat{\tau}_k := \tau_k \text{ et } \eta_k := 1,$$

sinon poser $\eta_k = \alpha_1^s$ où s est un nombre non négatif avec

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^r \tau_k)\|_F \leq \beta\tau_k \alpha_1^r, \quad r = 0, 1, \dots, s \text{ et}$$

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^{s+1} \tau_k)\|_F > \beta\tau_k \alpha_1^{s+1},$$

et poser

$$\widehat{\tau}_k := \eta_k \tau_k \text{ et } \widehat{W}^k := \begin{cases} W^k & \text{si } s=0 \\ W^k + \Delta W^k & \text{sinon.} \end{cases}$$

Etape 2: Pas Correcteur.

Soit $(\Delta \widehat{W}^k, \Delta \widehat{\tau}_k) = (\Delta \widehat{X}^k, \Delta \widehat{\lambda}^k, \Delta \widehat{S}^k, \Delta \widehat{\tau}_k)$ une solution de

$$\nabla \Theta(\widehat{W}^k, \widehat{\tau}_k) \begin{pmatrix} \Delta \widehat{W} \\ \Delta \widehat{\tau} \end{pmatrix} = -\Theta(\widehat{W}^k, \widehat{\tau}_k) + \begin{pmatrix} 0 \\ (1 - \hat{\sigma})\widehat{\tau}_k \end{pmatrix}. \quad (3.7)$$

Soit $\widehat{\eta}_k$ le maximum des nombres $1, \alpha_2, \alpha_2^2, \dots$, tel que

$$\|\phi(\widehat{X}^k + \widehat{\eta}_k \Delta \widehat{X}^k, \widehat{S}^k + \widehat{\eta}_k \Delta \widehat{S}^k, \widehat{\tau}_k + \widehat{\eta}_k \Delta \widehat{\tau}_k)\|_F \leq (1 - \hat{\sigma}\widehat{\eta}_k)\beta\widehat{\tau}_k \quad (3.8)$$

poser

$$W^{k+1} := \widehat{W}^k + \widehat{\eta}_k \Delta \widehat{W}^k, \tau_{k+1} := (1 - \widehat{\sigma} \widehat{\eta}_k) \tau_k, k \leftarrow k + 1,$$

et aller à l'Etape 1.

Il est aisé de constater que tous les itérés (X^k, λ^k, S^k) et $(\widehat{X}^k, \widehat{\lambda}^k, \widehat{S}^k)$ engendrés par l'algorithme ci-dessus sont admissibles pour les conditions d'optimalité (1.3) dans le sens que les deux équations suivantes

$$\begin{aligned} \sum_{i=1}^m \lambda_i^k A_i + S^k &= C, A_i \bullet X^k = b_i \quad (i = 1, \dots, m) \\ \text{et} \\ \sum_{i=1}^m \widehat{\lambda}_i^k A_i + \widehat{S}^k &= C, A_i \bullet \widehat{X}^k = b_i \quad (i = 1, \dots, m) \end{aligned} \tag{3.9}$$

sont satisfaites pour tout $k \in \mathbb{N}$.

Nous verrons par la suite que toutes les matrices X^k, S^k et $\widehat{X}^k, \widehat{S}^k$ sont automatiquement symétriques. Ceci constitue une différence avec les méthodes de points intérieurs. Ces dernières doivent, pour assurer la symétrie de leurs directions, symétriser les conditions de la trajectoire centrale. Mais il est important de souligner que les matrices de cet algorithme ne sont pas nécessairement définies positives ou semi-définies positives.

Le critère de terminaison utilisé dans l'Etape 1 de l'algorithme, i.e., $\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F = 0$ est justifié par les propositions 2.1 et 2.3 ($\phi(X, S) = 0 \Leftrightarrow X \succeq 0, S \succeq 0, XS = 0$). Combinant ceci avec notre discussion sur l'admissibilité des itérés, nous avons

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F = 0$$

ssi

$W^k + \Delta W^k$ est une solution des conditions d'optimalité.

Pour l'analyse théorique de l'algorithme, le critère de terminaison n'est jamais satisfait lorsque l'algorithme génère une suite infinie.

La règle de mise à jour de τ_{k+1} dans l'Etape 2 de l'algorithme peut être formulée de manière plus standard, i.e., $\tau_{k+1} = \widehat{\tau}_k + \widehat{\eta}_k \Delta \widehat{\tau}_k$. Cette observation résulte de la dernière équation du système linéaire (3.7) du pas Correcteur de l'algorithme. En effet, cette équation donne $\Delta \widehat{\tau}_k = \widehat{\sigma} \widehat{\tau}_k$.

Dans l'algorithme sus-cité, nous devons résoudre un système linéaire

d'équations dans le pas Prédicteur ainsi que dans le pas Correcteur. Cette résolution se fait en utilisant les deux matrices possibles. Nous entendons par là, les deux formulations possibles de $\nabla\Theta(W, \tau)$. Le coût est donc plus élevé que ce qui est habituellement fait dans le cas des méthodes de points intérieurs.

Malgré tout, les résultats de convergence obtenus restent vérifiés avec la modification suivante de l'algorithme : si le pas Prédicteur est accepté avec $\eta_k < 1$, nous passons outre le pas Correcteur. Nous posons donc $W^{k+1} := W^k + \Delta W^k$, $\tau_{k+1} := \eta_k \tau_k$, $k = k + 1$ et allons à l'Etape 1. L'algorithme ne doit plus résoudre qu'un système linéaire. Dans l'autre cas, les deux systèmes à résoudre possèdent la même matrice des coefficients.

Le but est de montrer que l'algorithme est bien défini. Il suffit de vérifier que les systèmes linéaires (3.6) du pas Prédicteur et (3.7) du pas Correcteur possèdent une solution unique. Pour cela, nous avons besoin de quelques propriétés de l'opérateur de Lyapunov.

Lemme 3.1 *Soient $A, B \in \mathcal{S}_{++}^{n \times n}$ et soient L_A et L_B les opérateurs de Lyapunov correspondant avec leur inverse L_A^{-1} et L_B^{-1} . Alors, les assertions suivantes sont équivalentes :*

- (a) L_A et L_B sont auto-adjoints.
- (b) L_A^{-1} et L_B^{-1} sont auto-adjoints.
- (c) $L_A \circ L_B$ et $L_B \circ L_A$ sont fortement monotones.
- (d) $L_A^{-1} \circ L_B$ et $L_B^{-1} \circ L_A$ sont fortement monotones.

Avant de donner la preuve du lemme, rappelons la définition : F est *fortement monotone* s'il existe $\rho \in \mathbb{R}_{++}$ tel que

$$\langle F(x) - F(y), x - y \rangle \geq \rho \|x - y\|^2 \quad \forall x, y \in \mathcal{S}.$$

Preuve du lemme 3.1 :

(a) L'opérateur L_A est auto-adjoint. En effet, soient $X, Y \in \mathcal{S}^{n \times n}$. Alors

$$\begin{aligned} L_A[X] \bullet Y &= \text{tr}(L_A[X]Y) = \text{tr}((AX + XA)Y) = \text{tr}(AXY) + \text{tr}(XAY) \\ &= \text{tr}(XYA) + \text{tr}(XAY) = \text{tr}(X(YA + AY)) \\ &= \text{tr}(XL_A[Y]) = X \bullet L_A[Y]. \end{aligned}$$

(b) L'opérateur L_A^{-1} est auto-adjoint. En effet, par la définition de l'inverse d'un opérateur et par (a), nous avons pour tous $X, S \in \mathcal{S}^{n \times n}$

$$\begin{aligned} L_A^{-1}[X] \bullet Y &= L_A^{-1}[X] \bullet L_A[L_A^{-1}[Y]] \\ &= L_A[L_A^{-1}[X]] \bullet L_A^{-1}[Y] \\ &= X \bullet L_A^{-1}[Y]. \end{aligned}$$

(c) En utilisant (a), nous obtenons, pour tout $X \in \mathcal{S}^{n \times n}$,

$$\begin{aligned} (L_A \circ L_B[X]) \bullet X &= L_B[X] \bullet L_A[X] = \text{tr}(L_B[X]L_A[X]) \\ &= \text{tr}((BX + XB)(AX + XA)) \\ &= \text{tr}(BXAX + XBAX + BXXA + XBXA) \\ &= \text{tr}(2BXAX + X^2(BA + AB)) \\ &= 2\text{tr}(BXAX) + \text{tr}(X^2(BA + AB)) \\ &= 2\|B^{1/2}XA^{1/2}\|_F^2 + \text{tr}(X(BA + AB)X). \end{aligned} \tag{3.10}$$

Puisque BA (AB) est semblable à $B^{1/2}AB^{1/2}$ ($A^{1/2}BA^{1/2}$) et que A, B sont symétriques définies positives, il suit que BA et AB ont leurs valeurs propres réelles et positives. D'où la matrice symétrique $BA + AB$ est définie positive. Par conséquent, la matrice $X(BA + AB)X$ est semi-définie positive et vérifie

$$\text{tr}(X(BA + AB)X) \geq 0$$

pour tout $X \in \mathcal{S}^{n \times n}$. De plus, comme la fonction $X \mapsto \|B^{1/2}XA^{1/2}\|_F$ définit une norme et que toutes les normes sont équivalentes en dimension finie, il existe une constante $\mu > 0$ telle que

$$\|B^{1/2}XA^{1/2}\|_F \geq \mu\|X\|_F \quad \forall X \in \mathcal{S}^{n \times n}. \tag{3.11}$$

Combinant (3.10) et (3.11), nous obtenons :

$$(L_A \circ L_B[X]) \bullet X \geq 2\|B^{1/2}XA^{1/2}\|_F^2 \geq 2\mu^2\|X\|_F^2$$

i.e., $L_A \circ L_B$ est fortement monotone sur $\mathcal{S}^{n \times n}$.

Pour voir que $L_B \circ L_A$ est également fortement monotone, il suffit d'intervertir le rôle de A et B .

(d) Puisque L_A est auto-adjoint par (a), nous obtenons, pour chaque $X \in \mathcal{S}^{n \times n}$,

$$\begin{aligned} (L_A^{-1} \circ L_B[X]) \bullet X &= (L_A^{-1} \circ L_B \circ L_A[Y]) \bullet L_A[Y] \\ &= (L_B \circ L_A[Y]) \bullet Y \end{aligned}$$

en posant $Y := L_A^{-1}[X]$. Cependant, $L_B \circ L_A$ est fortement monotone par (c). D'où (d) suit de (c). □

Si nous parvenons à montrer que la fonction linéaire $\nabla\Theta(X, \lambda, S, \tau)$ est inversible, nous pourrions en déduire que les systèmes linéaires (3.6) et (3.7) de l'algorithme ont une solution unique. Pour atteindre notre objectif, nous introduisons l'hypothèse suivante :

Hypothèse 3.1 *Les matrices A_i ($i = 1, \dots, m$) sont linéairement indépendantes.*

En utilisant le lemme 3.1 et cette hypothèse, nous arrivons à démontrer que $\nabla\Theta(X, \lambda, S, \tau)$ est une bijection. Notons que, dans ce cas, la direction $(\Delta X^k, \Delta \lambda^k, \Delta S^k, \Delta \tau_k)$ du pas Prédicteur ainsi que celle du pas Correcteur $(\Delta \hat{X}^k, \Delta \hat{\lambda}^k, \Delta \hat{S}^k, \Delta \hat{\tau}_k)$ seront bien définies.

Proposition 3.1 *Supposons que l'hypothèse 3.1 est satisfaite. Alors, la fonction linéaire $\nabla\Theta(X, \lambda, S, \tau)$, avec ϕ définie par (2.12) ou (2.13), est bijective pour tout $(X, \lambda, S, \tau) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}_{++}$.*

Preuve :

Soient $\phi(X, S, \tau) := X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2}$ et

$(X, \lambda, S, \tau) \in \mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}_{++}$ fixé. Comme $\nabla\Theta(X, \lambda, S, \tau)$ est une fonction linéaire de l'espace vectoriel $\mathcal{S}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}^{n \times n} \times \mathbb{R}$ de dimension finie dans lui-même, nous vérifions seulement que cette fonction

est injective. Il est suffisant de montrer que le système
 $\nabla\Theta(X, \lambda, S, \tau)(\Delta X, \Delta\lambda, \Delta S, \Delta\tau) = (0, 0, 0, 0)$
ou, de manière équivalente, que le système

$$\sum_{i=1}^m \Delta\lambda_i A_i + \Delta S = 0 \quad (3.12)$$

$$A_i \bullet \Delta X = 0 \quad (i = 1, \dots, m) \quad (3.13)$$

$$\nabla\phi(X, S, \tau)(\Delta X, \Delta S, \Delta\tau) = 0 \quad (3.14)$$

$$\Delta\tau = 0 \quad (3.15)$$

a $(\Delta X, \Delta\lambda, \Delta S, \Delta\tau) = (0, 0, 0, 0)$ comme seule solution.

De la dernière équation de ce système, nous obtenons $\Delta\tau = 0$.

En posant $E := (X^2 + S^2 + 2\tau^2 I)^{1/2}$, grâce au théorème 2.2 et à l'équation (3.14), nous avons

$$\Delta X + \Delta S - L_E^{-1}[X\Delta X + \Delta X X + S\Delta S + \Delta S S] = 0.$$

L'application de L_E des deux côtés de l'équation et la réorganisation des termes nous donnent

$$L_{E-X}[\Delta X] + L_{E-S}[\Delta S] = 0.$$

Comme $E - S \succ 0$, l'inverse L_{E-S}^{-1} existe et nous obtenons :

$$L_{E-S}^{-1} \circ L_{E-X}[\Delta X] + \Delta S = 0. \quad (3.16)$$

Par (3.12) et (3.13) et en prenant le produit scalaire avec ΔX , l'équation (3.16) devient

$$\begin{aligned} 0 &= L_{E-S}^{-1} \circ L_{E-X}[\Delta X] \bullet \Delta X - \sum_{i=1}^m \Delta\lambda_i A_i \bullet \Delta X \\ &= L_{E-S}^{-1} \circ L_{E-X}[\Delta X] \bullet \Delta X. \end{aligned} \quad (3.17)$$

Comme $E - X \succ 0$ et $E - S \succ 0$, il suit du lemme 3.1 (d) que l'opérateur $L_{E-S}^{-1} \circ L_{E-X}$ est fortement monotone. Par conséquent, (3.17) implique immédiatement que $\Delta X = 0$.

En effet, par la définition de fortement convexe, nous déduisons de $0 \geq \rho \|\Delta X\|^2$ où $\rho > 0$ que $\Delta X = 0$. Nous tirons ensuite de (3.16) que $\Delta S = 0$. Enfin, l'hypothèse d'indépendance linéaire des matrices et (3.12) entraînent que $\Delta\lambda = 0$.

Pour la fonction ϕ définie par (2.13), la démarche est analogue.

□

Dans cette démonstration, nous avons supposé que la matrice $E - S$ était définie positive. Montrons-le grâce au lemme suivant.

Lemme 3.2 Soient $E \in \mathcal{S}_{++}^{n \times n}$ et $S \in \mathcal{S}^{n \times n}$. Si $E^2 - S^2 \in \mathcal{S}_{++}^{n \times n}$ alors $E - S \in \mathcal{S}_{++}^{n \times n}$.

Preuve :

Fixons $E \in \mathcal{S}_{++}^{n \times n}$ et $S \in \mathcal{S}^{n \times n}$ tels que $E^2 - S^2 \in \mathcal{S}_{++}^{n \times n}$.

Si nous prouvons que

$$E - |S| \in \mathcal{S}_{++}^{n \times n}$$

où $|S| = (S^2)^{1/2}$, alors $E - S$ sera symétrique défini positif car $|S| - S$ est symétrique semi-définie positive (cfr. [11, page 182]).

Supposons, par l'absurde, que $E - |S| \notin \mathcal{S}_{++}^{n \times n}$. Alors

$$\exists v \in \mathbb{R}^n, v \neq 0 \text{ et } \exists \lambda \in (-\infty, 0] \text{ tels que } (E - |S|)v = \lambda v.$$

Sachant que l'égalité suivante

$$(E + |S|)(E - |S|) + (E - |S|)(E + |S|) = 2E^2 - 2S^2 \in \mathcal{S}_{++}^{n \times n},$$

est vraie, nous obtenons

$$0 < v^T [(E + |S|)(E - |S|) + (E - |S|)(E + |S|)]v.$$

Or $(E - |S|)v = \lambda v$. D'où

$$0 < \lambda v^T (E + |S|)v + \lambda v^T (E + |S|)v = 2\lambda v^T (E + |S|)v \leq 0.$$

La dernière inégalité vient du fait que $(E + |S|)$ est défini positif, $v \neq 0$ et $\lambda \leq 0$. Nous arrivons ainsi à une contradiction. La supposition est donc fausse et la thèse est satisfaite. □

En se basant sur les précédents résultats, nous montrons que l'algorithme est bien défini sous l'hypothèse 3.1.

Théorème 3.1 L'algorithme est bien défini sous l'hypothèse 3.1. De plus, les itérés $W^k = (X^k, \lambda^k, S^k)$, τ_k , $\widehat{W}^k = (\widehat{X}^k, \widehat{\lambda}^k, \widehat{S}^k)$ et $\widehat{\tau}_k$ appartiennent au voisinage $\mathcal{N}(\beta)$.

Preuve :

Grâce à la proposition 3.1, il nous suffit de montrer que les deux stratégies de backtracking de l'algorithme, i.e., celle de l'Etape 1 et celle de l'Etape 2, sont bien définies.

1) Considérons la première stratégie. Si nous supposons que l'algorithme engendre une suite infinie, le critère de terminaison dans l'Etape 1 n'est pas vérifié. D'où,

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F > 0$$

pour n'importe quel $k \in \mathbb{N}$. Puisque la fonction $\tau \rightarrow \|\phi(X, S, \cdot)\|_F$ est continue, la stratégie de backtracking de l'Etape 1 se termine en un nombre fini de boucles internes.

En effet, prenons $s = 0$ et regardons la stratégie de backtracking.

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^r \tau_k)\|_F \leq \beta \tau_k \alpha_1^r, r = 0, 1, \dots, s,$$

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^{s+1} \tau_k)\|_F > \beta \tau_k \alpha_1^{s+1},$$

Dans ce cas, $\alpha_1^s = 1$ et la première inégalité est automatiquement satisfaite. S'il n'existe pas de s tel que

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^{s+1} \tau_k)\|_F > \beta \tau_k \alpha_1^{s+1},$$

soit vrai, nous aurions

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^{s+1} \tau_k)\|_F \leq \beta \tau_k \alpha_1^{s+1},$$

et en passant à la limite

$$\begin{aligned} \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^{s+1} \tau_k)\|_F &> 0 \\ \text{et} \\ \beta \tau_k \alpha_1^{s+1} &\rightarrow 0. \end{aligned}$$

Ce qui est impossible. D'où, les itérés \widehat{W}^k et $\widehat{\tau}^k$ sont bien définis. De plus, ils remplissent la condition de voisinage. Prouvons-le.

Soit $W^0 = (X^0, \lambda^0, S^0)$ tel que

$$\|\phi(X^0, S^0, \tau_0)\|_F \leq \beta \tau_0.$$

Si

$$\|\phi(X^0 + \Delta X^0, S^0 + \Delta S^0, \tau_0)\|_F > \beta \tau_0,$$

la mise à jour est

$$\widehat{W}^0 := W^0, \widehat{\tau}_0 := \tau_0 \text{ et } \eta_0 := 1.$$

Comme les valeurs prises par les différents paramètres sont inchangées, la condition est toujours satisfaite.

Par ailleurs, si l'inégalité

$$\|\phi(X^0 + \Delta X^0, S^0 + \Delta S^0, \tau_0)\|_F \leq \beta \tau_0$$

est vraie, nous calculons $\eta_k = \alpha_1^s$ où s est un nombre non négatif avec

$$\|\phi(X^0 + \Delta X^0, S^0 + \Delta S^0, \alpha_1^r \tau_0)\|_F \leq \beta \tau_0 \alpha_1^r, \quad r = 0, 1, \dots, s,$$

$$\|\phi(X^0 + \Delta X^0, S^0 + \Delta S^0, \alpha_1^{s+1} \tau_0)\|_F > \beta \tau_0 \alpha_1^{s+1}.$$

La mise à jour est alors

$$\hat{\tau}_0 := \eta_k \tau_0 \text{ et } \widehat{W}^0 := \begin{cases} W^0 & \text{si } s=0 \\ W^0 + \Delta W^0 & \text{sinon.} \end{cases}$$

En se souvenant de (3.9), l'itéré vérifie donc la condition de voisinage.

La preuve est semblable pour les autres itérés $k \geq 1$.

2) Considérons à présent la longueur du pas de l'Etape 2 et définissons la fonction $\psi(X, S, \tau) := \|\phi(X, S, \tau)\|_F$. Alors, les règles standards de calcul différentiel impliquent

$$\begin{aligned} & \psi'(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)(\Delta \hat{X}^k, \Delta \hat{S}^k, \Delta \hat{\tau}_k) \\ &= \frac{\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k) \bullet \nabla \phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)(\Delta \hat{X}^k, \Delta \hat{S}^k, \Delta \hat{\tau}_k)}{\|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F} \\ &= -\|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F \end{aligned} \quad (3.18)$$

Supposant que le calcul de la longueur du pas ne se termine pas en un nombre fini de boucles internes à l'Etape 2, nous obtenons, pour tout $t \in \mathbb{N}$,

$$\|\phi(\hat{X}^k + \alpha_2^t \Delta \hat{X}^k, \hat{S}^k + \alpha_2^t \Delta \hat{S}^k, \hat{\tau}_k + \alpha_2^t \Delta \hat{\tau}_k)\|_F > (1 - \hat{\sigma} \alpha_2^t) \beta \hat{\tau}_k.$$

L'inégalité $\beta \hat{\tau}_k \geq \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F$ étant satisfaite car les itérés appartiennent au voisinage, nous en déduisons

$$\|\phi(\hat{X}^k + \alpha_2^t \Delta \hat{X}^k, \hat{S}^k + \alpha_2^t \Delta \hat{S}^k, \hat{\tau}_k + \alpha_2^t \Delta \hat{\tau}_k)\|_F > (1 - \hat{\sigma} \alpha_2^t) \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F$$

ou de manière équivalente

$$\begin{aligned} & \frac{\|\phi(\hat{X}^k + \alpha_2^t \Delta \hat{X}^k, \hat{S}^k + \alpha_2^t \Delta \hat{S}^k, \hat{\tau}_k + \alpha_2^t \Delta \hat{\tau}_k)\|_F - \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F}{\alpha_2^t} \\ & > -\hat{\sigma} \|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F \quad \forall t \in \mathbb{N}. \end{aligned}$$

Passant à la limite $t \rightarrow \infty$ et utilisant (3.18), nous avons

$$\begin{aligned} -\|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F &= \psi'(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)(\Delta\hat{X}^k, \Delta\hat{S}^k, \Delta\hat{\tau}_k) \\ &\geq -\hat{\sigma}\|\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k)\|_F. \end{aligned}$$

D'où, $\phi(\hat{X}^k, \hat{S}^k, \hat{\tau}_k) = 0$ car $\hat{\sigma} \in (0, 1)$. Donc, quand $t \rightarrow \infty$,

$$\|\phi(\hat{X}^k + \alpha_2^t \Delta\hat{X}^k, \hat{S}^k + \alpha_2^t \Delta\hat{S}^k, \hat{\tau}_k + \alpha_2^t \Delta\hat{\tau}_k)\|_F \rightarrow 0.$$

Par ailleurs, $(1 - \hat{\sigma}\alpha_2^t)\beta\hat{\tau}_k \rightarrow \beta\hat{\tau}_k$ lorsque $t \rightarrow \infty$ car $\alpha_2^t \rightarrow 0$. Or, par hypothèse, nous avons, pour tout $t \in \mathbb{N}$,

$$\|\phi(\hat{X}^k + \alpha_2^t \Delta\hat{X}^k, \hat{S}^k + \alpha_2^t \Delta\hat{S}^k, \hat{\tau}_k + \alpha_2^t \Delta\hat{\tau}_k)\|_F > (1 - \hat{\sigma}\alpha_2^t)\beta\hat{\tau}_k.$$

Ce qui n'est pas possible. La recherche linéaire de l'Etape 2 est par conséquent bien définie.

L'appartenance des itérés au voisinage $\mathcal{N}(\beta)$ est une conséquence directe des règles de mise à jour de l'algorithme.

□

Chapitre 4

Convergence, Calcul du point de départ et Critère d'arrêt.

4.1 Convergence globale.

Remarquons que les règles de mise à jour du paramètre τ dans notre algorithme assurent la décroissance monotone de ce paramètre et sa positivité à chaque itération.

Théorème 4.1 *Si la suite $\{W^k\} = \{(X^k, \lambda^k, S^k)\}$ engendrée par l'algorithme possède un point d'accumulation, alors la suite $\{\tau_k\}$ converge vers zéro. En particulier, chaque point d'accumulation de cette suite $\{W^k\}$ est une solution des conditions d'optimalité (1.3).*

Preuve :

1) Puisque la suite décroît de façon monotone et est bornée inférieurement par zéro, elle converge vers $\tau_* \geq 0$. Si $\tau_* = 0$, c'est fini. Supposons $\tau_* > 0$. Alors, les règles de mise à jour à l'Etape 1 de l'algorithme donnent

$$\widehat{W}^k = W^k, \widehat{\tau}_k = \tau_k \text{ et } \eta_k = 1 \quad (4.1)$$

pour tout $k \in \mathbb{N}$ suffisamment grand. Si nécessaire, nous supposons, sans perdre de généralité, que (4.1) est vérifiée pour tout $k \in \mathbb{N}$.

Nous avons $\tau_* > 0$ donc nous ne pouvons avoir la mise à jour $\tau_k = \eta_k \tau_k$. En effet, $\eta_k \rightarrow 0$ au fur et à mesure que s augmente car $\eta_k = \alpha_1^s$ et nous aurions alors $\tau_* = 0$, ce qui est contraire à l'hypothèse faite sur τ_* .

La règle de mise à jour de l'Etape 2 de l'algorithme donne

$$\tau_k = \tau_0 \prod_{j=0}^{k-1} (1 - \widehat{\sigma} \widehat{\eta}_j).$$

Comme, par hypothèse, $\tau_k \rightarrow \tau_* > 0$, il suit que $\lim_{k \rightarrow +\infty} \widehat{\eta}_k = 0$.

La taille du pas $\widehat{\rho}_k := \widehat{\eta}_k / \alpha_2$ ne satisfait pas le critère de recherche linéaire (3.8) du pas Correcteur pour tout $k \in \mathbb{N}$ suffisamment grand.

En effet, $\widehat{\eta}_k$ est le premier des nombres $1, \alpha_2, \alpha_2^2, \dots$, tel que l'équation (3.8) soit vérifiée. Cette équation n'est donc pas satisfaite pour le précédent, i.e., $\widehat{\eta}_k / \alpha_2 = \widehat{\rho}_k$. Nous avons donc

$$\|\phi(\widehat{X}^k + \widehat{\rho}_k \Delta \widehat{X}^k, \widehat{S}^k + \widehat{\rho}_k \Delta \widehat{S}^k, \widehat{\tau}_k + \widehat{\rho}_k \Delta \widehat{\tau}_k)\|_F > (1 - \widehat{\sigma} \widehat{\rho}_k) \beta \widehat{\tau}_k. \quad (4.2)$$

pour tous ces $k \in \mathbb{N}$.

Maintenant, soient $W^* = (X^*, \lambda^*, S^*)$ un point d'accumulation de la suite $\{W^k\}$ et $\{W^k\}_K$ une sous-suite convergeant vers W^* . Puisque $\tau_* > 0$, par (4.1) et la proposition 3.1, la suite correspondante $\{(\Delta \widehat{W}^k, \Delta \widehat{\tau}_k)\}_K$ converge vers $(\Delta \widehat{W}^*, \Delta \widehat{\tau}_*) = (\Delta \widehat{X}^*, \Delta \widehat{\lambda}^*, \Delta \widehat{S}^*, \Delta \widehat{\tau}_*)$ où $(\Delta \widehat{W}^*, \Delta \widehat{\tau}_*)$ est solution du système linéaire

$$\nabla \Theta(W^*, \tau_*) \begin{pmatrix} \Delta \widehat{W} \\ \Delta \widehat{\tau} \end{pmatrix} = -\Theta(W^*, \tau_*) + \begin{pmatrix} 0 \\ (1 - \widehat{\sigma}) \tau_* \end{pmatrix}. \quad (4.3)$$

En particulier, la suite $\{(\Delta \widehat{W}^k, \Delta \widehat{\tau}_k)\}_K$ est bornée. En utilisant $\{\widehat{\rho}_k\}_K \rightarrow 0$ et en prenant la limite $k \rightarrow \infty$ sur le sous-ensemble K , nous obtenons de (4.1), (4.2) et de la continuité de la fonction $\phi(\cdot, \cdot, \cdot)$

$$\|\phi(X^*, S^*, \tau_*)\|_F \geq \beta \tau_* > 0. \quad (4.4)$$

D'autre part, nous obtenons de (4.1), de (4.2), de la définition de voisinage $\mathcal{N}(\beta)$ et du théorème 3.1

$$\begin{aligned} & \|\phi(\widehat{X}^k + \widehat{\rho}_k \Delta \widehat{X}^k, \widehat{S}^k + \widehat{\rho}_k \Delta \widehat{S}^k, \widehat{\tau}_k + \widehat{\rho}_k \Delta \widehat{\tau}_k)\|_F \\ & > (1 - \widehat{\sigma} \widehat{\rho}_k) \beta \widehat{\tau}_k \\ & = (1 - \widehat{\sigma} \widehat{\rho}_k) \beta \tau_k \\ & \geq (1 - \widehat{\sigma} \widehat{\rho}_k) \|\phi(X^k, S^k, \tau_k)\|_F \end{aligned}$$

pour tout $k \in \mathbb{N}$ suffisamment grand. En tenant compte de (4.1), l'inégalité devient

$$\frac{\|\phi(\hat{X}^k + \hat{\rho}_k \Delta \hat{X}^k, \hat{S}^k + \hat{\rho}_k \Delta \hat{S}^k, \hat{\tau}_k + \hat{\rho}_k \Delta \hat{\tau}_k)\|_F - \|\phi(X^k, S^k, \tau_k)\|_F}{\hat{\rho}_k} > -\hat{\sigma} \|\phi(X^k, S^k, \tau_k)\|_F.$$

La fonction $\psi(X, S, \tau) := \|\phi(X, S, \tau)\|_F$ étant continûment différentiable en (X^*, S^*, τ_*) car $\tau_* > 0$, le passage à la limite dans l'inégalité précédente nous donne grâce à (3.18)

$$-\|\phi(X^*, S^*, \tau_*)\|_F \geq -\hat{\sigma} \|\phi(X^*, S^*, \tau_*)\|_F \quad (4.5)$$

Comme $\hat{\sigma} \in (0, 1)$, nous avons alors $\|\phi(X^*, S^*, \tau_*)\|_F = 0$ qui est en contradiction avec (4.4).

2) Soient $W^* = (X^*, \lambda^*, S^*)$ un point d'accumulation de la suite $\{W^k\} = \{(X^k, \lambda^k, S^k)\}$ déterminée par l'algorithme et $\{W^k\}_K$ une sous-suite convergeant vers W^* . La première partie du théorème prouve que

$$\lim_{k \rightarrow \infty} \tau_k = 0.$$

Comme tous les itérés appartiennent au voisinage $\mathcal{N}(\beta)$, par le théorème 3.1, nous en déduisons

$$\|\phi(X^*, S^*, 0)\|_F = \lim_{k \in K} \|\phi(X^k, S^k, \tau_k)\|_F \leq \lim_{k \in K} \beta \tau_k = 0$$

L'admissibilité des itérés et les propositions 2.1 et 2.3 impliquent que

$$W^* \text{ est solution des conditions d'optimalité (1.3).}$$

□

La conséquence principale de ce théorème est la convergence globale de l'algorithme. Remarquons que le résultat de convergence globale dépend seulement du pas Correcteur. Le choix précis du pas Prédicteur dans notre algorithme n'a aucune influence sur ce résultat.

4.2 Convergence locale.

L'objectif à atteindre : prouver que la suite $\{\tau_k\}$ converge superlinéairement vers zéro.

Définition : La suite $\{\tau_k\}$ converge superlinéairement vers 0 s'il existe une suite $c_k \rightarrow 0$ telle que $\|\tau_{k+1}\| \leq c_k \|\tau_k\|$ pour k suffisamment grand.

Comme la convergence locale dépend de certaines propriétés du pas Prédicteur de notre algorithme, énonçons l'hypothèse suivante :

Hypothèse 4.1 La suite $\{\tau_k\}$ engendrée par l'algorithme converge vers 0 et

$$\left\| \begin{pmatrix} \Delta W^k \\ \Delta \tau_k \end{pmatrix} \right\| = O(\tau_k) \quad (4.6)$$

où $(\Delta W^k, \Delta \tau_k)$ est la direction calculée dans le système (3.6) du pas Prédicteur.

Pour justifier cette hypothèse, nous constatons d'abord que le théorème 4.1 est une condition suffisante pour que la suite $\{\tau_k\}$ converge vers zéro. Pour comprendre l'équation (4.6), nous supposons que la suite des opérateurs inverses $\nabla\Theta(W^k, \tau_k)^{-1}$ est bornée lorsque $k \rightarrow \infty$. Ainsi, nous obtenons du système linéaire (3.6) de notre algorithme que (4.6) est satisfait, à condition que le membre de droite du système (3.6) soit de l'ordre de $O(\tau_k)$. En effet, quand $\nabla\Theta^{-1}$ est borné, (3.6) nous donne

$$\left\| \begin{pmatrix} \Delta W^k \\ \Delta \tau_k \end{pmatrix} \right\| \leq c \underbrace{\| -\Theta(W^k, \tau_k) \|}_{O(\tau_k)}.$$

Nous arrivons au résultat souhaité.

Il reste à prouver que le membre de droite du système linéaire (3.6) de l'algorithme est de l'ordre de $O(\tau_k)$. De part l'admissibilité des itérés et l'appartenance de ces derniers au voisinage $\mathcal{N}(\beta)$, nous avons

$$\begin{aligned} \| \Theta(W^k, \tau_k) \| &= \sqrt{\| \phi(X^k, S^k, \tau_k) \|_F^2 + \tau_k^2} \\ &\leq \| \phi(X^k, S^k, \tau_k) \|_F + \tau_k \\ &\leq \beta \tau_k + \tau_k = O(\tau_k). \end{aligned}$$

Cette relation sera aussi satisfaite si nous remplaçons le membre de droite du système linéaire (3.6) par $-\Theta(W^k, 0)$. En effet, le corollaire 2.1 et le

théorème 3.1 impliquent

$$\begin{aligned}
|||\Theta(W^k, 0)||| &= \|\phi(X^k, S^k, 0)\|_F \\
&\leq \|\phi(X^k, S^k, \tau_k) - \phi(X^k, S^k, 0)\|_F + \|\phi(X^k, S^k, \tau_k)\|_F \\
&\leq \kappa\sqrt{n}\tau_k + \beta\tau_k \\
&= O(\tau_k)
\end{aligned}$$

où $\kappa > 0$ est la constante donnée par le lemme 2.2.

En particulier, toutes les propriétés de convergences locale et globale de l'algorithme restent vérifiées si nous utilisons cette modification du membre de droite du système (3.6).

Afin d'énoncer une condition suffisante pour que l'hypothèse 4.1 soit satisfaite, introduisons l'hypothèse suivante :

Hypothèse 4.2 Soit (X^*, λ^*, S^*) une solution des conditions d'optimalité (1.3) telle que

- (a) (Complémentarité stricte) $X^* + S^* \succ 0$,
- (b) (Non dégénérescence) Pour tout $(\Delta X, \Delta \lambda, \Delta S)$ satisfaisant

$$\sum_{i=1}^m \Delta \lambda_i A_i + \Delta S = 0 \text{ et } A_i \bullet \Delta X = 0 \quad (i = 1, \dots, m),$$

l'implication suivante est vérifiée :

$$X^* \Delta S + \Delta X S^* = 0 \Rightarrow (\Delta X, \Delta S) = (0, 0).$$

Par le prochain théorème, l'hypothèse 4.1 sera vérifiée sous les hypothèses 3.1 et 4.2 si les itérés (X^k, λ^k, S^k) engendrés par l'algorithme convergent vers une solution (X^*, λ^*, S^*) satisfaisant les deux conditions de l'hypothèse 4.2.

Théorème 4.2 Supposons que les hypothèses 3.1 et 4.2 sont vérifiées en une solution (X^*, λ^*, S^*) des conditions d'optimalité.
Alors la fonction linéaire $\nabla \Theta(X^*, \lambda^*, S^*, 0)$ est bijective.

Preuve :

Considérons le cas où ϕ est donnée par (2.12). Définissons $E := ((X^*)^2 +$

$(S^*)^2)^{1/2}$. Par la supposition de complémentarité stricte, il est aisé de constater que E est une matrice définie positive.

En effet, nous savons que $X^* + S^* \succ 0$ et X^*, S^* est solution de (1.3). D'où $\phi(X^*, S^*) = 0$, i.e., $X^* + S^* = ((X^*)^2 + (S^*)^2)^{1/2}$ et donc E est définie positive.

Le théorème 2.2 nous permet alors de dire que Θ est continûment différentiable en $(X^*, \lambda^*, S^*, 0)$.

Pour voir que $\nabla\Theta(X^*, \lambda^*, S^*, 0)$ est bijective, nous devons seulement montrer qu'elle est injective. Pour cela, considérons l'équation

$$\nabla\Theta(X^*, \lambda^*, S^*, 0)(\Delta X, \Delta\lambda, \Delta S, \Delta\tau) = (0, 0, 0, 0)$$

et prouvons que $(\Delta X, \Delta\lambda, \Delta S, \Delta\tau) = (0, 0, 0, 0)$ est la seule solution. La dernière équation donne

$$\Delta\tau = 0. \quad (4.7)$$

En tenant compte de ceci et en utilisant le théorème 2.2, nous pouvons réécrire les trois premières équations comme suit

$$\sum_{i=1}^m \Delta\lambda_i A_i + \Delta S = 0, \quad (4.8)$$

$$A_i \bullet \Delta X = 0 \quad (i = 1, \dots, m), \quad (4.9)$$

$$\Delta X + \Delta S - L_E^{-1}[X^* \Delta X + \Delta X X^* + S^* \Delta S + \Delta S S^*] = 0. \quad (4.10)$$

Par la preuve de la proposition 3.1, l'équation (4.10) devient

$$L_{E-X^*}[\Delta X] + L_{E-S^*}[\Delta S] = 0. \quad (4.11)$$

Comme (X^*, λ^*, S^*) est une solution strictement complémentaire des conditions d'optimalité, le produit $X^* S^* = 0$ et donc X^* et S^* commutent. Ces deux matrices peuvent être diagonalisées simultanément par une transformation orthogonale. Nous pouvons trouver une seule matrice orthogonale $Q \in \mathbb{R}^{n \times n}$ et des matrices diagonales D_X et $D_S \in \mathbb{R}^{n \times n}$ telles que

$$X^* = Q^T D_X Q \text{ et } S^* = Q^T D_S Q. \quad (4.12)$$

Par (4.12), un simple calcul donne

$$E - X^* = S^* \text{ et } E - S^* = X^*. \quad (4.13)$$

L'équation (4.11) peut être mise sous la forme suivante :

$$S^* \Delta X + \Delta X S^* + X^* \Delta S + \Delta S X^* = 0.$$

Combinant (4.8), (4.9), l'hypothèse (4.2) et le lemme 6.2 de [14], nous avons

$$(\Delta X, \Delta S) = (0, 0).$$

Comme les matrices A_i sont linéairement indépendantes par l'hypothèse 3.1, il suit de (4.8) que

$$\Delta \lambda = 0.$$

Le cas où ϕ est définie par (2.13) est analogue. □

Nous soulignons que ce théorème fournit seulement une condition suffisante pour que l'hypothèse 4.1 soit satisfaite.

Débutons l'analyse du comportement local de l'algorithme avec un résultat technique.

Lemme 4.1 *Supposons l'hypothèse 4.1 satisfaite. Alors*

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k + \Delta \tau_k)\|_F = o(\tau_k).$$

Preuve :

Par le système linéaire (3.6) de l'algorithme, nous avons

$$\nabla \phi(X^k, S^k, \tau_k)(\Delta X^k, \Delta S^k, \Delta \tau_k) = -\phi(X^k, S^k, \tau_k).$$

Appliquer le théorème de la valeur moyenne sous forme intégrale donne

$$\begin{aligned} & \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k + \Delta \tau_k)\|_F \\ &= \left\| \int_0^1 \nabla \phi(X^k + \eta \Delta X^k, S^k + \eta \Delta S^k, \tau_k + \eta \Delta \tau_k) \begin{pmatrix} \Delta X^k \\ \Delta S^k \\ \Delta \tau_k \end{pmatrix} d\eta \right. \\ & \quad \left. + \phi(X^k, S^k, \tau_k) \right\|_F \\ &= \left\| \int_0^1 \left[\nabla \phi(X^k + \eta \Delta X^k, S^k + \eta \Delta S^k, \tau_k + \eta \Delta \tau_k) - \nabla \phi(X^k, S^k, \tau_k) \right] \right. \\ & \quad \left. \begin{pmatrix} \Delta X^k \\ \Delta S^k \\ \Delta \tau_k \end{pmatrix} d\eta \right\|_F \end{aligned}$$

$$\begin{aligned}
&\leq \int_0^1 \left\| \left[\nabla \phi(X^k + \eta \Delta X^k, S^k + \eta \Delta S^k, \tau_k + \eta \Delta \tau_k) - \nabla \phi(X^k, S^k, \tau_k) \right] \right. \\
&\quad \left. \begin{pmatrix} \Delta X^k \\ \Delta S^k \\ \Delta \tau_k \end{pmatrix} \right\|_F d\eta \\
&= o \left(\left\| \begin{pmatrix} \Delta X^k \\ \Delta S^k \\ \Delta \tau_k \end{pmatrix} \right\| \right)
\end{aligned}$$

où la dernière égalité provient du fait que ϕ est continûment différentiable. Utiliser l'hypothèse 4.1 et $W^k = (X^k, S^k, \tau_k)$ donne

$$\|(\Delta W^k)\| \leq c\tau_k.$$

La thèse suit. □

Grâce au prochain résultat, nous serons à même de prouver la convergence locale superlinéaire de la suite $\{\tau_k\}$

Lemme 4.2 *Supposons l'hypothèse 4.1 satisfaite. Soit β une constante vérifiant l'inégalité $\beta > \kappa\sqrt{n}$ où κ désigne la constante du lemme 2.2. Alors la suite $\{\eta_k\}$ définie dans l'Etape 2 de l'algorithme converge vers zéro.*

Preuve :

Soit $\epsilon > 0$ arbitraire. Par le système linéaire (3.6) de notre algorithme, nous savons que $\Delta\tau_k = -\tau_k$. Du lemme 4.1, nous obtenons

$$\begin{aligned}
&\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F \\
&= \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \tau_k + \Delta\tau_k)\|_F \\
&= o(\tau_k).
\end{aligned}$$

D'où, il existe un index $K_\epsilon \in \mathbb{N}$ tel que

$$\|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F \leq \epsilon\tau_k \quad \forall k \geq K_\epsilon.$$

Alors, $\forall \eta > 0, \forall k \geq K_\epsilon$, nous avons

$$\begin{aligned}
& \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \eta \tau_k)\|_F \\
& \leq \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F \\
& + \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \eta \tau_k) - \phi(X^k + \Delta X^k, S^k + \Delta S^k, 0)\|_F \\
& \leq \epsilon \tau_k + \kappa \eta \tau_k \sqrt{n},
\end{aligned}$$

où la dernière inégalité suit du corollaire 2.1.

Comme $\epsilon \tau_k + \kappa \sqrt{n} \eta \tau_k \leq \beta \tau_k$ est vérifié pour tout $\eta \geq \frac{\epsilon}{\beta - \kappa \sqrt{n}}$, la définition de η_k nous montre que $\eta_k \alpha_1$ ne satisfait pas cette inégalité, i.e.,

$$\eta_k < \frac{\epsilon}{(\beta - \kappa \sqrt{n}) \alpha_1}. \quad (4.14)$$

En effet, dans l'algorithme, $\eta_k = \alpha_1^s$ et

$$\begin{aligned}
& \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^s \tau_k)\|_F \leq \beta \tau_k \alpha_1^s \\
& \text{et } \|\phi(X^k + \Delta X^k, S^k + \Delta S^k, \alpha_1^{s+1} \tau_k)\|_F > \beta \tau_k \alpha_1^{s+1}.
\end{aligned}$$

Puise $\beta - \kappa \sqrt{n} > 0$ par hypothèse et $\epsilon > 0$ choisi arbitrairement, (4.14) implique que

$$\eta_k \rightarrow 0.$$

□

Etablissons le résultat principal de la convergence locale de l'algorithme.

Théorème 4.3 *Sous l'hypothèse 4.1, nous avons $\tau_{k+1} = o(\tau_k)$, i.e., le paramètre de régularisation τ_k converge localement superlinéairement vers zéro.*

Preuve :

Tenant compte du lemme 4.2 et de la définition de τ_{k+1} et $\hat{\tau}_k$ dans l'algorithme, nous arrivons à

$$\tau_{k+1} \leq \hat{\tau}_k = \eta_k \tau_k = o(\tau_k),$$

i.e.,

$$\tau_k \rightarrow 0 \text{ superlinéairement.}$$

□

Quelques remarques concernant ce théorème :

Le théorème reste satisfait si nous remplaçons le membre de droite du système linéaire (3.6) du pas Prédicteur par $-\Theta(W^k, 0)$. Cela suit de l'analyse faite en [3]. Grâce à [3], nous pouvons également montrer que $\nabla\phi$ est localement Lipschitzienne. D'où, par cette propriété, nous sommes à même de prouver que le taux de convergence est quadratique dans le théorème 4.3. En empruntant un résultat à [15], nous pouvons établir que Θ est fortement semi-régulière, au moins si Θ est définie par la fonction minimum. Combiner ce fait avec l'hypothèse de régularité forte 4.2(b) à une solution des conditions d'optimalité entraîne que la convergence locale de notre méthode est quadratique.

4.3 Calcul du point de départ et Critère d'arrêt.

Afin de décrire la façon dont nous calculons le point de départ (X^0, λ^0, S^0) , appelons le triplet (X, λ, S) **admissible** pour les conditions d'optimalité (1.3) s'il satisfait les équations linéaires

$$\sum_{i=1}^m \lambda_i A_i + S = C \quad \text{admissibilité du dual}$$

et

$$A_i \bullet X = b_i \quad i = 1, \dots, m \quad \text{admissibilité du primal.}$$

Bien entendu, le point de départ (X^0, λ^0, S^0) doit être admissible en ce sens. Mais, notons que nous ne demandons pas que $X \succeq 0$ ou $S \succeq 0$ pour un tel triplet admissible.

A cette fin, nous définissons une matrice symétrique $\mathcal{A} \in \mathbb{R}^{m \times m}$ par $\mathcal{A}_{ij} = A_i \bullet A_j$ pour $i, j = 1, \dots, m$. Nous résolvons ensuite le système linéaire $\mathcal{A}y = b$ pour obtenir $y^0 \in \mathbb{R}^m$. Nous définissons alors

$$X^0 = \sum_{i=1}^m y_i^0 A_i$$

et calculons λ^0 comme la solution du système

$$\mathcal{A}\lambda = (A_1 \bullet C, \dots, A_m \bullet C)^T.$$

Finalement, nous posons

$$S^0 = C - \sum_{i=1}^m \lambda_i^0 A_i$$

et le point de départ (X^0, λ^0, S^0) sera admissible.

Ayant calculé le point de départ, les paramètres restant dans l'algorithme sont initialisés par

$$\tau_0 = \frac{\|\phi(X^0, S^0, 0)\|_F}{5}$$

et

$$\beta = \max \left\{ 2.1\sqrt{n}, 1.5 \frac{\|\phi(X, S, \tau_0)\|}{\tau_0} \right\}.$$

Critère d'arrêt.

Nous terminons l'itération si

$$\tau_k/n < 10^{-6}$$

(nous paramétrisons les conditions de la trajectoire centrale par τ^2) et si la mesure d'admissibilité

$$\max \left\{ \frac{\|[b_i - A_i \bullet X^k]_{i=1}^m\|_2}{\max\{1, \|b\|_2\}}, \frac{\|C - S^k - \sum_{i=1}^m \lambda_i^k A_i\|_F}{\max\{1, \|C\|_2\}} \right\}$$

est plus petite que 10^{-10} . La raison de diviser par τ_k par n est basée sur le fait que

$$\|\phi(X^k, S^k, 0)\|_F = O(\tau_k).$$

Comme nous voulons avoir la norme $\|\phi(X^k, S^k, 0)\|_F$ petite, il semble raisonnable de terminer si τ_k est petit. Par ailleurs, obtenir la norme $\|\phi(X^k, S^k, 0)\|_F$ petite devient moins aisé lorsque les dimensions de X^k et S^k sont grandes car nous utilisons la norme $\|\cdot\|_F$. Pour rendre le critère d'arrêt plus ou moins indépendant de la dimension de X^k et S^k , nous prendrons les règles sus-mentionnées.

Chapitre 5

Quelques aspects pratiques d'une méthode de type Newton pour les SPDs.

Par la proposition 3.1, nous savons que les systèmes de Newton (3.6) du pas Prédicteur et (3.7) du pas Correcteur ont une seule solution. Le calcul des solutions de ces systèmes est le principal effort de calcul quand nous appliquons l'algorithme aux problèmes de programmation semi-définie. Pour notre analyse, nous considérerons le système (3.6) où le membre de droite est remplacé par $-\Theta(W^k, 0)$. Nous allons fournir deux stratégies pour la solution de ces systèmes de Newton. Nous donnerons aussi quelques propriétés de ces systèmes.

Nous avons vu que Θ pouvait être définie à l'aide de deux fonctions. La première, la fonction minimum, la deuxième, la fonction de Fischer-Burmeister. Nous baserons notre analyse sur la fonction minimum. Nous conclurons en regardant si les résultats obtenus sont satisfaits pour la fonction de Fischer-Burmeister.

5.1 Formulation de Lyapunov des systèmes de Newton.

Notre but est de montrer que les systèmes de Newton peuvent être résolus en utilisant l'inverse de la fonction de Lyapunov. Nous le ferons seulement pour le système (3.6) modifié.

Introduisons les résidus

$$R_C := C - \sum_{j=1}^m \lambda_j A_j - S,$$

$$r_{b,i} := b_i - A_i \bullet X \quad (i = 1, \dots, m),$$

$$r_b := (r_{b,1}, \dots, r_{b,m})^T.$$

Alors, le système de Newton devient

$$\sum_{j=1}^m \Delta \lambda_j A_j + \Delta S = R_C, \quad (5.1)$$

$$A_i \bullet \Delta X = r_{b,i} \quad (i = 1, \dots, m), \quad (5.2)$$

$$\nabla \phi(X, S, \tau)(\Delta X, \Delta S, \Delta \tau) = -\phi(X, S, 0), \quad (5.3)$$

$$\Delta \tau = 0, \quad (5.4)$$

où nous avons omis l'indice d'itération. Par le théorème 2.2 et par (5.4), il suit de (5.3) que

$$\begin{aligned} & \Delta X + \Delta S - L_E^{-1}[(X - S)(\Delta X - \Delta S) + (\Delta X - \Delta S)(X - S)] \\ & = -\phi(X, S, 0), \end{aligned} \quad (5.5)$$

où

$$E := ((X - S)^2 + 4\tau^2 I)^{1/2}$$

est la matrice symétrique définie positive du théorème 2.2. Si nous appliquons l'opérateur de Lyapunov L_E des deux côtés de (5.5), nous obtenons

$$\begin{aligned} & L_E[\Delta X] + L_E[\Delta S] - (X - S)(\Delta X - \Delta S) - (\Delta X - \Delta S)(X - S) \\ & = -L_E[\phi(X, S, 0)]. \end{aligned}$$

En arrangeant les termes, nous avons

$$L_{E-(X-S)}[\Delta X] + L_{E+(X-S)}[\Delta S] = -L_E[\phi(X, S, 0)].$$

Posant

$$A_E := E - (X - S) \text{ et } B_E := E + (X - S), \quad (5.6)$$

l'équation peut être réécrite sous la forme :

$$L_{A_E}[\Delta X] + L_{B_E}[\Delta S] = -L_E[\phi(X, S, 0)]. \quad (5.7)$$

Comme A_E de (5.6) est symétrique définie positive (B_E l'est aussi), nous obtenons

$$\Delta X = -L_{A_E}^{-1} [L_{B_E}[\Delta S] + L_E[\phi(X, S, 0)]] . \quad (5.8)$$

En substituant ΔS de (5.1) et en ordonnant les termes, nous avons

$$\Delta X = \sum_{j=1}^m \Delta \lambda_j L_{A_E}^{-1} [L_{B_E}[A_j]] - L_{A_E}^{-1} [L_{B_E}[R_C] + L_E[\phi(X, S, 0)]] .$$

En prenant les produits scalaires avec A_i ($i = 1, \dots, m$) et en considérant (5.2), nous arrivons à

$$\begin{aligned} & \sum_{j=1}^m \Delta \lambda_j L_{A_E}^{-1} [L_{B_E}[A_j]] \bullet A_i \\ &= A_i \bullet \Delta X + L_{A_E}^{-1} [L_{B_E}[R_C] + L_E[\phi(X, S, 0)]] \bullet A_i \\ &= r_{b,i} + L_{A_E}^{-1} [L_{B_E}[R_C] + L_E[\phi(X, S, 0)]] \bullet A_i \end{aligned} \quad (5.9)$$

pour $i = 1, \dots, m$. Cette expression pourra être transformée en lui appliquant le prochain résultat qui établit quelques propriétés de l'opérateur de Lyapunov.

Lemme 5.1 Soient $A, B \in \mathcal{S}_{++}^{n \times n}$, L_A, L_B leur opérateur de Lyapunov correspondant et L_A^{-1}, L_B^{-1} leur inverse respectif. Alors, les énoncés suivants sont satisfaits :

1. L_A, L_B, L_A^{-1} et L_B^{-1} sont auto-adjoints.
2. $L_A^{-1} \circ L_B$ et $L_B^{-1} \circ L_A$ sont fortement monotones.
3. Si A et B commutent, alors $L_A \circ L_B = L_B \circ L_A$.

Preuve : Nous ne montrons que 3.

Soient A et B deux matrices qui commutent, i.e., $AB = BA$. Alors, nous

obtenons

$$\begin{aligned}
(L_A \circ L_B)[X] &= L_A[L_B[X]] \\
&= L_A[BX + XB] \\
&= A(BX + XB) + (BX + XB)A \\
&= ABX + AXB + BXA + XBA \\
&= BAX + BXA + AXB + XAB \\
&= B(AX + XA) + (AX + XA)B \\
&= (L_B \circ L_A)[X]
\end{aligned}$$

pour chaque matrice $X \in \mathcal{S}^{n \times n}$. Par conséquent la thèse. □

Grâce au lemme 5.1 où $A := A_E$ et $B := B_E$ (A_E et B_E sont les matrices de (5.6)), (5.9) est équivalent à

$$\sum_{j=1}^m \Delta \lambda_j L_{B_E}[A_j] \bullet L_{A_E}^{-1}[A_i] = r_{b,i} + (L_{B_E}[R_C] + L_E[\phi(X, S, 0)]) \bullet L_{A_E}^{-1}[A_i], \quad (5.10)$$

pour $i = 1, \dots, m$. C'est une équation linéaire en les variables $\Delta \lambda \in \mathbb{R}^m$ où les éléments de la matrice des coefficients $M \in \mathbb{R}^{m \times m}$ sont définis par

$$m_{ij} := L_{B_E}[A_j] \bullet L_{A_E}^{-1}[A_i] \quad (i, j = 1, \dots, m). \quad (5.11)$$

Une fois ce système résolu, nous obtenons ΔS de (5.1) et ΔS est évidemment symétrique puisque R_C et toutes les matrices A_i le sont.

En vue de (5.8), ΔX est calculé comme une solution de l'équation de Lyapunov avec un membre de droite symétrique. Par conséquent, ΔX l'est aussi. (cfr. [11, th 2.2.3])

Considérons la matrice $M = (m_{ij})$ définie par (5.11). A première vue, cette matrice semble très générale. Notre objectif est de montrer que la matrice M est symétrique définie positive. La symétrie de cette matrice nous permettra de ne pas calculer tous ses éléments. Le fait qu'elle soit définie positive nous autorise à appliquer la factorisation de Cholesky (ou la méthode du gradient conjugué) afin de résoudre le système linéaire (5.10) pour $\Delta \lambda$.

Théorème 5.1 *Supposons que l'hypothèse 3.1 est vérifiée. Alors, la matrice $M \in \mathbb{R}^{m \times m}$ avec les m_{ij} donnés par (5.11) est définie positive.*

Preuve : Soit $d \in \mathbb{R}^m$ un vecteur donné arbitrairement. Alors

$$\begin{aligned}
d^T M d &= \sum_{i,j=1}^m d_i d_j m_{ij} \\
&= \sum_{i,j=1}^m d_i d_j \left(L_{B_E}[A_j] \bullet L_{A_E}^{-1}[A_i] \right) \\
&= \sum L_{B_E}[d_j A_j] \bullet L_{A_E}^{-1}[d_i A_i] \\
&= \sum_{i,j=1}^m \left(L_{A_E}^{-1} \circ L_{B_E} \right) [d_j A_j] \bullet [d_i A_i] \\
&= \sum_{j=1}^m \left(L_{A_E}^{-1} \circ L_{B_E} \right) [d_j A_j] \bullet \left(\sum_{i=1}^m d_i A_i \right) \\
&= \left(L_{A_E}^{-1} \circ L_{B_E} \right) \left[\sum_{j=1}^m d_j A_j \right] \bullet \left[\sum_{i=1}^m d_i A_i \right] \geq 0
\end{aligned}$$

par les points 1 et 2 du lemme 5.1. De plus, le point 2 du lemme 5.1 implique que l'égalité est seulement satisfaite pour $\sum_{i=1}^m d_i A_i = 0$.
L'indépendance linéaire des matrices A_1, \dots, A_m entraîne $d_i = 0 \ i = 1, \dots, m$. □

Maintenant, nous allons prouver que M est symétrique. Pour cela, rappelons le résultat suivant :

Proposition 5.1 Soient $A, B \in \mathcal{S}^{n \times n}$. Alors, les énoncés suivants sont équivalents :

1. A et B commutent.
2. A et B ont une décomposition spectrale simultanée, i.e., il existe une matrice orthogonale $Q \in \mathbb{R}^{n \times n}$ et deux matrices $D_A, D_B \in \mathbb{R}^{n \times n}$ telles que $A = Q^T D_A Q$ et $B = Q^T D_B Q$.

Pour pouvoir appliquer le point 3 du lemme 5.1 montrons que les deux matrices A_E et B_E de (5.6) commutent.

Lemme 5.2 Les deux matrices A_E et B_E de (5.6) commutent.

Preuve :

Nous devons démontrer que $A_E B_E = B_E A_E$. Pour ce faire, grâce à la proposition 5.1, prouvons que les deux matrices A_E et B_E possèdent une décomposition spectrale simultanée.

Soit

$$X - S = Q^T D Q$$

la décomposition spectrale de la matrice symétrique $X - S$. Alors,

$$(X - S)^2 = Q^T D^2 Q$$

est la décomposition spectrale de la matrice $(X - S)^2$ et nous obtenons par conséquent

$$(X - S)^2 + 4\tau^2 I = Q^T (D^2 + 4\tau^2 I) Q$$

et

$$E = ((X - S)^2 + 4\tau^2 I)^{1/2} = Q^T (D^2 + 4\tau^2 I)^{1/2} Q.$$

D'où

$$A_E = E - (X - S) = Q^T ((D^2 + 4\tau^2 I)^{1/2} - D) Q$$

et

$$B_E = E + (X - S) = Q^T ((D^2 + 4\tau^2 I)^{1/2} + D) Q.$$

montrent que A_E et B_E ont une décomposition spectrale simultanée.

□

Nous sommes donc en mesure d'établir la symétrie de la matrice M .

Théorème 5.2 *Soit $M = (m_{ij})$ la matrice dont les éléments sont définis par 5.11. Alors M est symétrique.*

Preuve :

Selon le lemme 5.2, les deux matrices symétriques définies positives A_E et B_E de (5.6) commutent. Le point 3 du lemme 5.1 implique que

$$L_{A_E} \circ L_{B_E} = L_{B_E} \circ L_{A_E}$$

ou

$$L_{A_E}^{-1} \circ L_{B_E} = L_{B_E} \circ L_{A_E}^{-1}. \quad (5.12)$$

En effet, $L_{A_E}^{-1} \circ L_{B_E} = L_{B_E} \circ L_{A_E}^{-1}$. Par composition à droite avec L_{A_E} nous avons $L_{A_E} \circ L_{A_E}^{-1} \circ L_{B_E} = L_{A_E} \circ L_{B_E} \circ L_{A_E}^{-1}$ ou $L_{B_E} = L_{A_E} \circ L_{B_E} \circ L_{A_E}^{-1}$. En composant à gauche avec L_{A_E} , nous pouvons conclure.

Pour montrer que la matrice M est symétrique, nous devons avoir $m_{ij} = m_{ji}$. Grâce au point 1 du lemme 5.1, c'est équivalent à

$$\begin{aligned} m_{ij} = m_{ji} &\Leftrightarrow L_{B_E}[A_j] \bullet L_{A_E}^{-1}[A_i] = L_{B_E}[A_i] \bullet L_{A_E}^{-1}[A_j] \\ &\Leftrightarrow (L_{A_E}^{-1} \circ L_{B_E})[A_j] \bullet A_i = A_i \bullet (L_{B_E} \circ L_{A_E}^{-1})[A_j] \\ &\Leftrightarrow (L_{A_E}^{-1} \circ L_{B_E})[A_j] \bullet A_i = (L_{B_E} \circ L_{A_E}^{-1})[A_j] \bullet A_i. \end{aligned}$$

La dernière égalité est satisfaite par (5.12). D'où la matrice M est symétrique. \square

5.2 Formulation Matrice-Vecteur des systèmes de Newton.

Nous allons exprimer les systèmes de Newton en des produits Matrice Vecteur. Pour y parvenir, nous devons mettre les matrices sous forme de vecteurs. (voir Annexe B)

Considérons les équations (5.1) à (5.4) de la section précédente. Les deux premières équations peuvent être formulées de la façon suivante grâce à la référence [16]:

$$\mathcal{A}^T \Delta \lambda + svec(\Delta S) = svec(R_C) \quad (5.13)$$

et

$$\mathcal{A} \text{vec}(\Delta X) = r_b \quad (5.14)$$

où

$$\mathcal{A} := (\text{vec}(A_1), \dots, \text{vec}(A_m))^T \in \mathbb{R}^{m \times \frac{n(n+1)}{2}}. \quad (5.15)$$

Il reste la troisième équation. Nous avons vu qu'elle pouvait se mettre sous la forme suivante :

$$A_E \Delta X + \Delta X A_E + B_E \Delta S + \Delta S B_E = -L_E[\phi(X, S, 0)]$$

où A_E et B_E sont les deux matrices symétriques définies positives de (5.6). En appliquant $\frac{1}{2} \text{vec}$ des deux côtés de l'équation, nous avons

$$\begin{aligned} \frac{1}{2} \text{vec}(A_E \Delta X + \Delta X A_E) + \frac{1}{2} \text{vec}(B_E \Delta S + \Delta S B_E) \\ = -\frac{1}{2} \text{vec}(L_E[\phi(X, S, 0)]). \end{aligned}$$

Par la définition (B.2) de vec , nous avons

$$\frac{1}{2} \text{vec}(A_E \Delta X + \Delta X A_E) = (I \otimes_S A_E) \text{vec}(\Delta X),$$

$$\frac{1}{2} \text{vec}(B_E \Delta S + \Delta S B_E) = (I \otimes_S B_E) \text{vec}(\Delta S).$$

Posant

$$\mathcal{E} := (I \otimes_S A_E) \text{ et } \mathcal{F} := (I \otimes_S B_E), \quad (5.16)$$

nous obtenons

$$\mathcal{E} \text{vec}(\Delta X) + \mathcal{F} \text{vec}(\Delta S) = -\frac{1}{2} \text{vec}(L_E[\phi(X, S, 0)]). \quad (5.17)$$

Résumons notre discussion et obtenons le résultat suivant comme conséquence de (5.13), (5.14) et (5.17).

Théorème 5.3 *Le triplet $(\Delta X, \Delta \lambda, \Delta S) \in S^{n \times n} \times \mathbb{R}^m \times S^{n \times n}$ vérifie le système de Newton (3.6) avec le membre de droite modifié si et seulement si le vecteur $(\text{vec}(\Delta X), \Delta \lambda, \text{vec}(\Delta S))$ satisfait le système linéaire d'équations*

$$\begin{pmatrix} 0 & \mathcal{A}^T & I \\ \mathcal{A} & 0 & 0 \\ \mathcal{E} & 0 & \mathcal{F} \end{pmatrix} \begin{pmatrix} \text{vec}(\Delta X) \\ \Delta \lambda \\ \text{vec}(\Delta S) \end{pmatrix} = \begin{pmatrix} \text{vec}(R_C) \\ r_b \\ -\frac{1}{2} \text{vec}(L_E[\phi(X, S, 0)]) \end{pmatrix}. \quad (5.18)$$

Notre objectif est de montrer que le système linéaire (5.18) possède une solution unique et comment cette solution peut être calculée. Afin d'y parvenir, nous avons besoin du résultat suivant :

Lemme 5.3 *Les matrices \mathcal{E} et \mathcal{F} de (5.16) sont symétriques définies positives. De plus, la matrice $\mathcal{E}^{-1}\mathcal{F}$ est aussi symétrique définie positive.*

Preuve :

Nous savons déjà que les matrices A_E et B_E de (5.6) sont symétriques définies positives. D'où, il suit des définitions de \mathcal{E} et \mathcal{F} de (5.16) et des points 1, 3 et 4 du lemme B.1 que \mathcal{E} et \mathcal{F} sont également symétriques définies positives.

En plus, par le lemme 5.2, les matrices A_E et B_E commutent. En utilisant les points 1 et 3 du lemme B.1, nous obtenons

$$\begin{aligned}\mathcal{E}\mathcal{F} &= (I \otimes_S A_E)(I \otimes_S B_E) \\ &= \frac{1}{2}(I \otimes_S A_E B_E + B_E \otimes_S A_E) \\ &= \frac{1}{2}(I \otimes_S B_E A_E + A_E \otimes_S B_E) \\ &= (I \otimes_S B_E)(I \otimes_S A_E) \\ &= \mathcal{F}\mathcal{E}.\end{aligned}$$

Donc, \mathcal{E} et \mathcal{F} commutent. Nous en déduisons $\mathcal{F}\mathcal{E}^{-1} = \mathcal{E}^{-1}\mathcal{F}$. Alors,

$$(\mathcal{E}^{-1}\mathcal{F})^T = \mathcal{F}^T(\mathcal{E}^{-1})^T = \mathcal{F}\mathcal{E}^{-1} = \mathcal{E}^{-1}\mathcal{F},$$

i.e., la matrice $\mathcal{E}^{-1}\mathcal{F}$ est symétrique. De plus, $\mathcal{E}^{-1}\mathcal{F}$ est définie positive puisqu'elle est semblable à la matrice symétrique définie positive $\mathcal{F}^{1/2}\mathcal{E}^{-1}\mathcal{F}^{1/2}$, où $\mathcal{F}^{1/2}$ désigne la racine carrée définie positive symétrique de la matrice \mathcal{F} . □

Le théorème suivant découle du lemme 5.3 et du [16, Thm 3.1].

Théorème 5.4 *Supposons les matrices A_i linéairement indépendantes. Alors, la matrice du système (5.18) est non singulière.*

Montrons comment nos précédents résultats peuvent être utilisés pour résoudre le système linéaire (5.18). Écrivons ce système comme suit

$$\mathcal{A}^T \Delta \lambda + \text{svec}(\Delta S) = \text{svec}(R_C), \quad (5.19)$$

$$\mathcal{A} \text{svec}(\Delta X) = r_b, \quad (5.20)$$

$$\mathcal{E} \text{svec}(\Delta X) + \mathcal{F} \text{svec}(\Delta S) = -\frac{1}{2} \text{svec}(L_E[\phi(X, S, 0)]). \quad (5.21)$$

Comme \mathcal{E} est non singulière par le lemme 5.3, nous obtenons de (5.21)

$$\text{svec}(\Delta X) = -\mathcal{E}^{-1} \left(\mathcal{F} \text{svec}(\Delta S) + \frac{1}{2} \text{svec}(L_E[\phi(X, S, 0)]) \right) \quad (5.22)$$

et après substitution de $\text{svec}(\Delta S)$ tiré de (5.19),

$$\text{svec}(\Delta X) = -\mathcal{E}^{-1} \left[\mathcal{F} \text{svec}(R_C) - \mathcal{F} \mathcal{A}^T \Delta \lambda + \frac{1}{2} \text{svec}(L_E[\phi(X, S, 0)]) \right].$$

En prémultipliant par \mathcal{A} et en utilisant (5.20), nous arrivons à

$$\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^T \Delta \lambda = r_b + \mathcal{A} \mathcal{E}^{-1} \left(\mathcal{F} \text{svec}(R_C) + \frac{1}{2} \text{svec}(L_E[\phi(X, S, 0)]) \right). \quad (5.23)$$

La procédure pour résoudre le système linéaire (5.18) est la suivante :

D'abord, calculer $\Delta \lambda$ de (5.23). La matrice de ce système est symétrique définie positive par le lemme 5.3 et par le fait que A soit de rang plein (sous l'hypothèse 3.1).

Alors, nous obtenons $\text{svec}(\Delta X)$ et $\text{svec}(\Delta S)$ de (5.22) et (5.19) respectivement.

En effet, $\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^T$ est définie positive puisque pour tout $x \neq 0$ $A^T x \neq 0$ car \mathcal{A} est de rang plein et $y^T \mathcal{E}^{-1} \mathcal{F} y > 0$ car $\mathcal{E}^{-1} \mathcal{F}$ est définie positive et $y \neq 0$.

Intéressons-nous à la relation entre la matrice

$$A := \mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^T \in \mathbb{R}^{m \times m} \quad (5.24)$$

de (5.23) et la matrice M correspondante de la section précédente définie par (5.11). Nous prétendons que ces matrices sont équivalentes. Pour y arriver, calculons les entrées a_{ij} de A . En utilisant la définition de A , nous obtenons

$$a_{ij} = \text{svec}(A_i)^T \mathcal{E}^{-1} \mathcal{F} \text{svec}(A_j).$$

Il suit de (B.2) que

$$\mathcal{F} \text{svec}(A_j) = (I \otimes_S B_E) \text{svec}(A_j) = \frac{1}{2} \text{svec}(B_E A_j + A_j B_E) = \frac{1}{2} \text{svec}(L_{B_E}[A_j]).$$

D'où nous obtenons

$$a_{ij} = \frac{1}{2} \text{svec}(A_i)^T (I \otimes_S A_E)^{-1} \text{svec}(L_{B_E}[A_j]). \quad (5.25)$$

D'autre part

$$A_E L_{A_E}^{-1}[A_i] + L_{A_E}^{-1}[A_i] A_E = A_i$$

par définition de l'opérateur inverse de Lyapunov. Par conséquent, (B.2) nous fournit

$$\begin{aligned} \frac{1}{2} \text{svec}(A_i) &= \frac{1}{2} \text{svec}(A_E L_{A_E}^{-1}[A_i] + L_{A_E}^{-1}[A_i] A_E) \\ &= (I \otimes_S A_E) \text{svec}(L_{A_E}^{-1}[A_i]). \end{aligned}$$

D'où

$$\frac{1}{2} (I \otimes_S A_E)^{-1} \text{svec}(A_i) = \text{svec}(L_{A_E}^{-1}[A_i]).$$

Si nous remplaçons ceci dans (5.25) et tenons compte de (B.1), alors

$$a_{ij} = \text{svec}(L_{A_E}^{-1}[A_i])^T \text{svec}(L_{B_E}[A_j]).$$

Or $A \bullet B = \text{svec}(A)^T \text{svec}(B)$ et $\text{tr}[AB] = \text{tr}[BA] = A \bullet B$. D'où

$$\begin{aligned} a_{ij} &= \text{svec}(L_{B_E}[A_j])^T \text{svec}(L_{A_E}^{-1}[A_i]) \\ &= L_{B_E}[A_j] \bullet L_{A_E}^{-1}[A_i]. \end{aligned}$$

La dernière ligne est identique à m_{ij} de (5.11).

De notre discussion, nous pouvons déduire le théorème

Théorème 5.5 *Les deux matrices $M = (m_{ij})$ de (5.11) et $A = (a_{ij})$ de (5.24) sont identiques.*

5.3 Résultats pour la fonction de Fischer-Burmeister.

Envisageons à partir de maintenant Θ défini par la fonction régulière de Fischer-Burmeister, i.e.,

$$\phi(X, S, \tau) := X + S - (X^2 + S^2 + 2\tau^2 I)^{1/2}. \quad (5.26)$$

La question est : "Quels résultats restent satisfaits si nous remplaçons la fonction régulière minimum par la fonction régulière de Fischer-Burmeister?" Regardons la première section.

L'équation (5.3) se met sous la forme suivante si nous considérons la fonction ϕ définie par (5.26)

$$\Delta X + \Delta S - L_E^{-1}[X\Delta X + \Delta X X + S\Delta S + \Delta S S] = -\phi(X, S, 0),$$

où $E := (X^2 + S^2 + 2\tau^2 I)^{1/2}$. Après quelques manipulations algébriques, semblables à celles de la première section, nous obtenons

$$L_{A_E}[\Delta X] + L_{B_E}[\Delta S] = -L_E[\phi(X, S, 0)],$$

où

$$A_E := E - X \text{ et } B_E := E - S. \quad (5.27)$$

Avec A_E et B_E définis comme ceci, nous sommes en mesure de calculer $\Delta\lambda \in \mathbb{R}^m$ en résolvant le système linéaire (5.10) où la matrice $M = (m_{ij}) \in \mathbb{R}^{m \times m}$ et ces éléments m_{ij} sont définis par (5.11).

Les matrices A_E et B_E déterminées par (5.27) sont définies positives. Il suit donc du théorème 5.1 que la matrice M est encore définie positive.

Cependant, les matrices A_E et B_E de (5.27) ne commutent pas en général. Une conséquence est que l'approche suivie dans la première section pour démontrer la symétrie de la matrice M n'est plus valable dans le cas de la fonction de Fischer-Burmeister. Il s'avère que la matrice M n'est généralement pas symétrique. A cette fin, prenons un contre-exemple.

Choisissons $n = 3$, $m = 2$, $r = 1$ et ϕ définie par (5.26).

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Soient les itérés

$$X = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } S = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Par un simple calcul, nous avons

$$E = (X^2 + S^2 + 2\tau^2 I)^{1/2} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Nous obtenons par conséquent

$$A_E = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, B_E = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix}.$$

D'où

$$L_{B_E}[A_1] = \begin{pmatrix} 4 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, L_{B_E}[A_i] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

et

$$L_{A_E}^{-1}[A_1] = \begin{pmatrix} \frac{7}{24} & \frac{1}{12} & 0 \\ \frac{1}{12} & \frac{7}{24} & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ et } L_{A_E}^{-1}[A_2] = \begin{pmatrix} \frac{1}{24} & \frac{1}{12} & 0 \\ \frac{1}{12} & \frac{7}{24} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

parce que $L_{A_E}^{-1}[A_1]A_E + A_EL_{A_E}^{-1}[A_1] = A_1$ et de façon semblable pour $L_{A_E}^{-1}[A_2]$. Nous voyons que

$$M = \begin{pmatrix} \frac{7}{6} & \frac{1}{12} \\ \frac{1}{6} & \frac{7}{12} \end{pmatrix}$$

n'est pas symétrique.

Considérons la deuxième section avec ϕ définie par (5.26) et A_E et B_E par (5.27). Nous pouvons suivre tous les arguments de cette section jusqu'au lemme 5.3. En effet, la preuve de ce lemme est basée sur le fait que A_E et B_E commutent. D'où, nous sommes toujours en mesure de conclure que la matrice $\mathcal{E}^{-1}\mathcal{F}$ est définie positive mais rien n'assure que cette matrice est symétrique.

Le système correspondant (5.18) a encore une solution et la matrice $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^T$ de (5.23) est également définie positive. Le théorème 5.5 et l'exemple montrent que la matrice $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^T$ n'est plus nécessairement symétrique.

Si nous définissons ϕ par la fonction minimum, le coût sera moindre que si elle est déterminée par la fonction de Fischer-Burmeister. Comme la symétrie n'est pas toujours garantie avec la fonction de Fischer-Burmeister, nous ne pouvons faire appel à la factorisation de Cholesky ni à la méthode du gradient.

Deuxième partie

Le cas des systèmes linéaires.

Chapitre 1

Les méthodes régulières Prédicteur-Correcteur pour la solution des Programmes linéaires.

1.1 Introduction.

Nous utilisons une classe de méthodes de type régulières Prédicteur-Correcteur pour résoudre les programmes linéaires. Soit le programme linéaire sous sa forme *Primale* :

$$P \equiv \begin{cases} \min & c^T x \\ \text{s.c.} & Ax = b, \\ & x \geq 0 \end{cases} \quad (1.1)$$

où $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$ et $b \in \mathbb{R}^m$ sont les données du problème. Nous supposons, sans perdre de généralité, que la matrice A est de rang plein. Le problème (1.1) possède la forme *Duale* suivante :

$$D \equiv \begin{cases} \max & b^T \lambda \\ \text{s.c.} & A^T \lambda + s = c, \\ & s \geq 0, \end{cases} \quad (1.2)$$

où s est une variable d'écart.

Les premières méthodes utilisées pour résoudre ce type de problème sont les méthodes du simplexe. Elles énumèrent les points extrêmes des ensembles

admissibles de (1.1) et (1.2) de façon intelligente jusqu'à ce qu'une solution optimale soit trouvée. Depuis lors, une autre sorte de méthode est apparue, celle de points intérieurs. Elle essaye de résoudre un programme linéaire en se basant sur la solution des conditions d'optimalité, i.e.,

$$\begin{aligned} A^T \lambda + s &= c, \\ Ax &= b, \\ x_i &\geq 0, s_i \geq 0, x_i s_i = 0 \quad \forall i = 1, \dots, n. \end{aligned} \tag{1.3}$$

Comme ces conditions d'optimalité sont précisément les conditions de K.K.T. à la fois du Primal et du Dual, il suit que

(1.1) a une solution ssi (1.2) en a une ssi (1.3) en possède une .

Par cette équivalence, résoudre (1.3) revient à trouver la solution de (1.1). La méthode présentée ici prend en compte le système (1.3). L'idée est de transformer le système (1.3) en un système non linéaire d'équations

$$\Phi(x, \lambda, s) = 0. \tag{1.4}$$

Cependant, Φ étant non différentiable, nous allons introduire un paramètre de régularisation $\tau > 0$ et remplacer le système non régulier (1.4) par une suite de systèmes différentiables

$$\Phi_\tau(x, \lambda, s). \tag{1.5}$$

Dans notre discussion, nous combinons les méthodes de type régulières avec l'idée de Jacobien régulier. Les pas de type régulier sont utilisés pour contrôler la convergence globale. Par ailleurs, les pas Jacobiens réguliers garantissent une convergence locale rapide. Nous insistons sur le fait que les hypothèses considérées dans la théorie de convergence locale ne requièrent pas que l'ensemble des solutions de (1.3) soit un singleton.

Introduisons quelques notations. Les normes des vecteurs sont les normes euclidiennes. Un vecteur $\hat{w} = (\hat{x}, \hat{\lambda}, \hat{s})$ est dit strictement admissible pour les conditions d'optimalité s'il satisfait $A^T \hat{\lambda} + \hat{s} = c$, $A \hat{x} = b$ et $\hat{x}_i > 0$, $\hat{s}_i > 0$ pour tout $i = 1, \dots, n$.

1.2 Fonctions et trajectoires régulières.

Les conditions d'optimalité peuvent s'exprimer comme un système non linéaire et habituellement non régulier d'équations en faisant appel aux fonctions NCP.

Définition : une fonction $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}$ est dite fonction NCP si elle a la propriété

$$\varphi(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0 \text{ et } ab = 0,$$

i.e., une fonction NCP vaut zéro exactement sur ces deux demi-axes non négatifs.

Soit φ une fonction NCP arbitraire, définissons $\Phi: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ par

$$\Phi(w) := \Phi(x, \lambda, s) := \begin{pmatrix} A^T + s - c \\ Ax - b \\ \phi(x, s) \end{pmatrix}, \quad (1.6)$$

où

$$\phi(x, s) := (\varphi(x_1, s_1), \dots, \varphi(x_n, s_n))^T \in \mathbb{R}^n.$$

Alors, il suit directement de la définition d'une fonction NCP que le vecteur $w^* = (x^*, \lambda^*, s^*)$ satisfait les conditions (1.3) si et seulement si w^* est une solution du système

$$\Phi(w) = 0. \quad (1.7)$$

Exemples de fonctions NCP :

1. $\varphi(a, b) := 2 \min\{a, b\}$ est appelée fonction minimum.
2. $\varphi(a, b) := a + b - \sqrt{a^2 + b^2}$ est appelée fonction de Fischer-Burmeister.

Puise φ et, par conséquent Φ , sont non différentiables, il nous est impossible d'appliquer la méthode de Newton classique au système (1.7). D'où, l'idée d'approximer φ et Φ par des fonctions régulières adéquates φ_τ et Φ_τ dépendant d'un certain paramètre de régularisation $\tau > 0$.

Exemples :

1. $\varphi_\tau(a, b) := a + b - \sqrt{(a - b)^2 + 4\tau^2}$, la fonction régulière minimum,
2. $\varphi_\tau(a, b) := a + b - \sqrt{a^2 + b^2 + 2\tau^2}$, la fonction régulière de Fischer-Burmeister.

Le prochain résultat nous montre que ces fonctions sont de bonnes approximations de leur fonction NCP non différentiable correspondante.

Lemme 1.1 *Soit φ une des deux fonctions NCP. Alors, il existe une constante $c = c_\varphi > 0$ (indépendante de τ et de (a, b)) telle que*

$$|\varphi(a, b) - \varphi_\tau(a, b)| \leq c\tau$$

pour tout $(a, b) \in \mathbb{R}^2$ et pour tout $\tau > 0$.

Preuve : Dans le cas de la fonction minimum, nous obtenons

$$\begin{aligned} |\varphi(a, b) - \varphi_\tau(a, b)| &= \left| a + b - \sqrt{(a-b)^2} - \left((a+b) - \sqrt{(a-b)^2 + 4\tau^2} \right) \right| \\ &= \sqrt{(a-b)^2 + 4\tau^2} - \sqrt{(a-b)^2} \\ &= \frac{(a-b)^2 + 4\tau^2 - (a-b)^2}{\sqrt{(a-b)^2 + 4\tau^2} + \sqrt{(a-b)^2}} \\ &\leq \frac{4\tau^2}{\sqrt{4\tau^2}} = 2\tau \end{aligned}$$

pour tout $(a, b) \in \mathbb{R}^2$ et pour tout $\tau > 0$, i.e., l'énoncé est satisfait avec $c = 2$.

Dans le cas de la fonction de Fischer-Burmeister, le résultat est vérifié pour $c = \sqrt{2}$.

□

Grâce à la définition des approximations φ_τ pour les fonctions NCP φ , nous sommes en mesure d'établir une approximation régulière de Φ de (1.6) par

$$\Phi_\tau(x, \lambda, s) := \Phi_\tau(w) := \begin{pmatrix} A^T \lambda + s - c \\ Ax - b \\ \phi_\tau(x, s) \end{pmatrix} \quad (1.8)$$

où $\phi_\tau(x, s) := (\varphi_\tau(x_1, s_1), \dots, \varphi_\tau(x_n, s_n))^T \in \mathbb{R}^n$.

Le résultat suivant nous montre la relation entre Φ et Φ_τ :

Lemme 1.2 Soit Φ définie ci-dessus.

Alors, il existe une constante $\kappa = \kappa_\Phi > 0$ (indépendante de τ et de w) telle que

$$\|\Phi(w) - \Phi_\tau(w)\| \leq \kappa\tau$$

pour tout $w = (x, \lambda, s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ et pour tout $\tau > 0$.

Preuve :

$$\begin{aligned} \|\Phi(w) - \Phi_\tau(w)\| &= \sqrt{|\phi(w) - \phi_\tau(w)|^2} \\ &= \sqrt{\sum_{i=1}^n |\phi(x_i, s_i) - \phi_\tau(x_i, s_i)|^2} \\ &\leq \sqrt{nc^2\tau^2} \\ &\leq c\tau\sqrt{n}. \end{aligned}$$

D'où $\kappa = c\sqrt{n}$.

□

Nous insistons sur le fait que la constante κ du lemme (1.2) est connue pour chaque fonction Φ .

Comme les approximations Φ_τ de Φ sont régulières, les matrices Jacobiennes $\Phi'_\tau(w)$ existent en chaque point w et pour chaque paramètre de régularisation $\tau > 0$. Un calcul élémentaire nous montre que

$$\Phi'_\tau(w) = \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ D_{a,\tau} & 0 & D_{b,\tau} \end{pmatrix} \quad (1.9)$$

où nous utilisons les abréviations

$$D_{a,\tau} := \text{diag}(\cdots, \frac{\partial \varphi_\tau}{\partial a}(x_i, s_i), \cdots), \quad (1.10)$$

$$D_{b,\tau} := \text{diag}(\cdots, \frac{\partial \varphi_\tau}{\partial b}(x_i, s_i), \cdots). \quad (1.11)$$

La proposition suivante établit que ces matrices Jacobiennes sont toujours non singulières. Par ailleurs, $\Phi'(w)$ qui est non pénalisée peut être singulière

même en des points différentiables, i.e., des points où la fonction pénalisée est régulière.

Proposition 1.1 *Soit Φ_τ dont l'expression est donnée en (1.8). Alors, les matrices Jacobiennes $\Phi'_\tau(w)$ sont non singulières pour chaque vecteur $w = (x, \lambda, s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ et chaque paramètre de régularisation $\tau > 0$.*

Preuve : Soit Φ_τ différentiable avec $\Phi'_\tau(x, \lambda, s)$ donnée par (1.9) et $D_{a, \tau} \in \mathbb{R}^{n \times n}$ et $D_{b, \tau} \in \mathbb{R}^{n \times n}$ les matrices diagonales de (1.10) et (1.11).

Comme $\frac{\partial \varphi_\tau}{\partial a}(a, b)$ et $\frac{\partial \varphi_\tau}{\partial b}(a, b) \in (0, 2)$ pour tout $(a, b) \in \mathbb{R}^2$ et par la définition de φ_τ , il suit que les matrices diagonales $D_{a, \tau}$ et $D_{b, \tau}$ sont définies positives. Pour montrer que Φ'_τ est non singulière, démontrons que

$$\Phi'_\tau(w)q = 0 \Rightarrow q = 0.$$

Soit $q = (q^{(1)}, q^{(2)}, q^{(3)}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ avec $\Phi'_\tau(w)q = 0$.

Alors, (1.9) implique

$$A^T q^{(2)} + q^{(3)} = 0, \quad (1.12)$$

$$Aq^{(1)} = 0, \quad (1.13)$$

$$D_{a, \tau} q^{(1)} + D_{b, \tau} q^{(3)} = 0. \quad (1.14)$$

Prémultipliant (1.12) par $(q^{(1)})^T$ et tenant compte de (1.13), nous en déduisons

$$(q^{(1)})^T q^{(3)} = 0. \quad (1.15)$$

D'autre part, en résolvant (1.14) pour $q^{(1)}$ et en le substituant dans (1.15), nous obtenons

$$(q^{(3)})^T D_{a, \tau}^{-1} D_{b, \tau} q^{(3)} = 0. \quad (1.16)$$

Du fait que $D_{a, \tau}$ et $D_{b, \tau}$ soient définies positives, nous déduisons de (1.16)

$$q^{(3)} = 0.$$

Par (1.14),

$$q^{(1)} = 0.$$

D'où, de (1.12), nous avons

$$q^{(2)} = 0$$

car la matrice A est de rang plein.

□

La prochaine proposition ainsi que ses corollaires garantiront l'existence de points d'accumulation quand nous utilisons une des fonctions pénalisées.

Proposition 1.2 *Soit ϕ la fonction pénalisée minimum ou de Fischer-Burmeister. Considérons l'ensemble*

$$\mathcal{L}_c := \{(x, \lambda, s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \mid A^T \lambda + s = c, Ax = b \text{ et } \|\Phi(w)\| \leq c\}$$

où $c \in \mathbb{R}$ une constante donnée et $w = (x, \lambda, s)$.

Supposons qu'il existe un point strictement admissible pour les conditions d'optimalité (1.3). Alors, l'ensemble \mathcal{L}_c est compact.

Preuve : cfr référence [8].

□

Corollaire 1.1 *Soit φ la fonction minimum ou de Fischer-Burmeister régulière. Considérons l'ensemble*

$$\mathcal{L}_{c, \tau} := \{(x, \lambda, s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \mid A^T \lambda + s = c, Ax = b, \|\Phi_\tau(w)\| \leq c\}$$

où $c \in \mathbb{R}$ constante donnée, $\tau \geq 0$ et $w = (x, \lambda, s)$.

Supposons qu'il existe un point strictement admissible pour les conditions d'optimalité. Alors, $\mathcal{L}_{c, \tau}$ est borné.

Preuve : cfr référence [8].

□

Corollaire 1.2 *Soient φ la fonction minimum ou de Fischer-Burmeister régulière et $\mathcal{L}_{c, \tau}$ l'ensemble défini dans le corollaire précédent.*

Alors, l'ensemble

$$\bigcup_{\tau \in [0, \hat{\tau}]} \mathcal{L}_{c, \tau}$$

est borné pour tout $c \in \mathbb{R}$, $\hat{\tau} \geq 0$.

Nous pouvons nous demander si l'équation régulière

$$\Phi_\tau(w) = 0 \quad (1.17)$$

possède une solution pour chaque $\tau > 0$.

Par la référence [12], nous avons que (1.17) est satisfaite ssi les conditions suivantes sont vérifiées :

$$\begin{aligned} A^T \lambda + s &= c, \\ Ax &= b, \\ x_i &> 0, s_i > 0, x_i s_i = \tau^2 \quad \forall i = 1, \dots, n. \end{aligned} \quad (1.18)$$

Ces conditions sont celles de la trajectoire centrale. Nous savons qu'elles ont une solution unique pour chaque $\tau > 0$ s'il existe un vecteur strictement admissible $\hat{w} = (\hat{x}, \hat{\lambda}, \hat{s})$ vérifiant les conditions d'optimalité.

Proposition 1.3 *Soit φ la fonction minimum ou de Fischer-Burmeister pénalisée. Supposons qu'il existe un vecteur strictement admissible pour les conditions d'optimalité. Alors, le système (1.17) possède une solution pour chaque $\tau > 0$.*

Preuve :

Soit $\tau > 0$ fixé.

Considérons le problème d'optimisation

$$\min \Psi_\tau(w) \text{ s.c. } A^T \lambda + s = c, Ax = b \quad (1.19)$$

où $\Psi_\tau : \mathbb{R}^n \rightarrow \mathbb{R}$ désigne la fonction

$$\Psi_\tau(w) := \frac{1}{2} \|\Phi_\tau(w)\|^2.$$

Comme A est de rang plein et l'ensemble $\mathcal{L}_{c, \tau}$ défini dans le corollaire 1.1 est non vide (pour $c > 0$ suffisamment grand) et compact. D'où, le problème d'optimisation (1.19) possède un minimum global $w_\tau = (x_\tau, \lambda_\tau, s_\tau)$. Nous montrons que w_τ est une solution du système non linéaire (1.17). Pour y parvenir, notons que (1.19) est un problème d'optimisation avec contraintes linéaires. D'où, il existe des multiplicateurs de Lagrange

$$\eta \in \mathbb{R}^m \text{ et } \mu \in \mathbb{R}^n$$

tels que (w_τ, η, μ) vérifie les conditions de K.K.T. de (1.19), i.e.,

$$\nabla \Psi_\tau(w_\tau) + \begin{pmatrix} A^T \eta \\ A\mu \\ \mu \end{pmatrix} = 0.$$

Puisque

$$\nabla \Psi_\tau(w_\tau) = (\Phi'_\tau(w_\tau))^T \Phi_\tau(w_\tau),$$

nous pouvons utiliser l'expression (1.9) pour le Jacobien afin d'écrire l'équation sus-citée comme

$$\begin{pmatrix} 0 & A^T & D_{a,\tau} \\ A & 0 & 0 \\ I & 0 & D_{b,\tau} \end{pmatrix} \begin{pmatrix} A^T \lambda_\tau + s_\tau - c \\ Ax_\tau - b \\ \phi_\tau(x_\tau, s_\tau) \end{pmatrix} + \begin{pmatrix} A^T \eta \\ A\mu \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

où les matrices diagonales $D_{a,\tau}$ et $D_{b,\tau}$ sont définies de façon similaire à (1.10) et (1.11). En plus, comme w_τ est admissible pour (1.19), le système précédent peut s'écrire

$$\begin{cases} D_{a,\tau} \phi_\tau(x_\tau, s_\tau) = -A^T \eta \\ A\mu = 0 \\ D_{b,\tau} \phi_\tau(x_\tau, s_\tau) + \mu = 0. \end{cases}$$

Multipliant la troisième équation par A , tenant compte de la deuxième équation et remplaçant $\phi_\tau(x_\tau, s_\tau)$ de la première équation, nous avons

$$AD_{a,\tau}^{-1} D_{b,\tau} A^T \eta = 0.$$

Puisque A est de rang plein et $D_{a,\tau}$ et $D_{b,\tau} \in \mathbb{R}^{n \times n}$ sont des matrices diagonales définies positives, $\eta = 0$. D'où, $\Phi_\tau(x_\tau, s_\tau) = 0$. Utilisant une fois de plus que w_τ satisfait les contraintes linéaires de (1.19), nous voyons que le vecteur w_τ est une solution du système non linéaire (1.17). \square

Cette proposition garantit l'existence d'une solution mais ne précise pas si elle est unique.

1.3 Méthode régulière Prédicteur-Correcteur.

L'idée de base de notre méthode est de trouver une solution des conditions d'optimalité en résolvant le système équivalent non linéaire d'équations

$$\Phi(w) = 0$$

où $\Phi(w)$ est défini par (1.6). Cependant, Φ étant non régulière et pouvant avoir des matrices Jacobiennes singulières, nous remplaçons le système $\Phi(w) = 0$ par le système différentiable

$$\Phi_\tau(w) = 0$$

auquel nous appliquerons la méthode de Newton.

Posons

$$\Psi_\tau(w) = \frac{1}{2} (\Phi_\tau(w))^T \Phi_\tau(w) = \frac{1}{2} \|\Phi_\tau(w)\|^2,$$

et énonçons notre algorithme.

Algorithme 1.

Etape 0: *Initialisation.*

Choisir $w^0 := (x^0, \lambda^0, s^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$, $\tau_0 > 0$, $\beta \geq \frac{\|\Phi_{\tau_0}(w^0)\|}{\tau_0}$.

Choisir $\rho, \sigma \in (0, 1)$, $\epsilon \geq 0$ et poser $k := 0$.

Etape 1: *Critère de terminaison.*

Si $\|\Phi(w^k)\| \leq \epsilon$: STOP.

Etape 2: *Pas Prédicteur.*

Calculer une solution $\Delta w^k = (\Delta x^k, \Delta \lambda^k, \Delta s^k) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ du système linéaire

$$\Phi'_{\tau_k}(w^k) \Delta w = -\Phi(w^k). \quad (1.20)$$

Si $\|\Phi(w^k + \Delta w^k)\| = 0$: STOP.

Sinon si

$$\|\Phi_{\tau_k}(w^k + \Delta w^k)\| > \beta \tau_k$$

alors poser $\hat{w}^k := w^k$ et $\hat{\tau}_k := \tau_k$

sinon calculer $t_k = \rho^{l_k}$ où l_k entier positif ou nul tel que

$$\|\Phi_{\rho^j \tau_k}(w^k + \Delta w^k)\| \leq \beta \rho^j \tau_k \quad \forall j = 0, 1, 2, \dots, l_k \text{ et}$$

$$\|\Phi_{\rho^{l_k+1} \tau_k}(w^k + \Delta w^k)\| > \beta \rho^{l_k+1} \tau_k,$$

et poser

$$\hat{\tau}_k := t_k \tau_k \text{ et } \hat{w}^k := \begin{cases} w^k & \text{si } l_k = 0 \\ w^k + \Delta w^k & \text{sinon.} \end{cases}$$

Etape 3 : Pas Correcteur.

Calculer la solution $\Delta\hat{w}^k = (\Delta\hat{x}^k, \Delta\hat{\lambda}^k, \Delta\hat{s}^k) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ du système linéaire

$$\Phi'_{\hat{\tau}_k}(\hat{w}^k)\Delta\hat{w} = -\Phi_{\hat{\tau}_k}(\hat{w}^k). \quad (1.21)$$

Soit $\hat{t}_k = \max\{\rho^l \mid l = 0, 1, 2, \dots\}$ tel que

$$\Psi_{\hat{\tau}_k}(\hat{w}^k + \hat{t}_k\Delta\hat{w}^k) \leq \Psi_{\hat{\tau}_k}(\hat{w}^k) + \hat{t}_k\sigma\nabla\Psi_{\hat{\tau}_k}(\hat{w}^k)\Delta\hat{w}^k \quad (1.22)$$

et poser

$$w^{k+1} := \hat{w}^k + \hat{t}_k\Delta\hat{w}^k.$$

Calculer $\gamma_k = \max\{\rho^l \mid l = 0, 1, 2, \dots\}$ tel que

$$\|\Phi_{(1-\gamma_k)\hat{\tau}_k}(w^{k+1})\| \leq \beta(1-\gamma_k)\hat{\tau}_k \quad (1.23)$$

et poser $\tau_{k+1} := (1-\gamma_k)\hat{\tau}_k$.

Etape 4 : Mise à jour.

Aller à l'Etape 1.

L'algorithme 1 résout deux systèmes linéaires à chaque itération, un dans le pas Prédicteur, un dans le pas Correcteur. Si le pas Prédicteur n'est pas réussi, i.e., $\|\Phi_{\tau_k}(w^k + \Delta w^k)\| > \beta\tau_k$, les coefficients des matrices de ces deux systèmes sont identiques. D'où, nous ne devons calculer qu'une seule matrice par itération.

Les propriétés de convergence globale vont dépendre du pas Correcteur tandis que la théorie locale est entièrement basée sur le pas Prédicteur.

Notons la différence majeure entre les systèmes linéaires (1.20) et (1.21). Le système linéaire (1.21) du pas Correcteur est juste le système de Newton appliqué à l'équation non linéaire $\Phi_{\hat{\tau}_k}(w) = 0$. Par contre, le membre de droite du système linéaire (1.20) du pas Prédicteur n'est pas perturbé, i.e., ne dépend pas du paramètre régulier. D'où, le système linéaire (1.20) ne peut être interprété comme une équation de Newton pure.

L'algorithme 1 tente d'engendrer une suite $\{w^k\}$ d'itérés et une suite $\{\tau_k\}$ de paramètres de régularisation vérifiant la relation

$$\|\Phi_{\tau_k}(w^k)\| \leq \beta\tau_k \quad (1.24)$$

pour tout $k \in \mathbb{N}$. De cette façon, nous garantissons que chaque w^k appartient à un certain voisinage de la trajectoire centrale. De plus, nous contrôlons la taille du voisinage grâce à notre paramètre régulier.

Faisons quelques commentaires sur l'algorithme 1 :

Après avoir résolu l'équation Jacobienne régulière (1.20), l'algorithme 1 vérifie si le pas $w^k + \Delta w^k$ appartient encore au voisinage (1.24) dans le sens

$$\|\Phi_{\tau_k}(w^k + \Delta w^k)\| \leq \beta \tau_k. \quad (1.25)$$

Si c'est vrai, nous posons

$$\hat{w}^k := w^k + \Delta w^k$$

et nous réduisons le paramètre τ_k autant que possible tout en s'assurant qu'une condition comme (1.24) soit toujours satisfaite à l'itéré intermédiaire \hat{w}^k et au nouveau paramètre $\hat{\tau}_k$.

D'autre part, si (1.25) n'est pas vérifiée, nous passons le pas Prédicteur et la mise à jour est la suivante : $\hat{w}^k := w^k$, $\hat{\tau}_k := \tau_k$. Cependant, il y a une exception qui peut devenir importante dans notre analyse de convergence globale : même si (1.25) est satisfaite, nous pouvons avoir quand $\hat{w}^k := w^k$ quand $l_k = 0$.

Le pas Correcteur calcule une nouvelle direction de Newton Δw^k dans (1.21), suivi par une recherche linéaire. Celle-ci vise à réduire la valeur de $\Psi_\tau(\cdot)$ de sorte que $w^{k+1} := \hat{w}^k + \hat{t}_k \Delta \hat{w}^k$ devienne le nouvel itéré où \hat{t}_k est la longueur du pas Correcteur. En plus, nous tentons de diminuer le paramètre régulier du pas Correcteur pour qu'une condition semblable à (1.24) soit vraie à l'itération $k + 1$.

Nous voulons montrer maintenant que l'algorithme 1 est bien défini. Pour ce faire, rappelons-nous que les deux systèmes linéaires (1.20) et (1.21) de l'algorithme 1 ont une solution unique. (C'est une conséquence de la proposition 1.1). Montrons que le calcul de $t_k > 0$ dans le pas Prédicteur est une procédure finie.

Proposition 1.4 *La longueur du pas $t_k > 0$ dans l'Etape 2 de l'algorithme 1 est définie de manière unique.*

Preuve :

Quand nous calculons la longueur du pas $t_k > 0$, nous nous trouvons dans

la situation où

$$\|\Phi(w^k + \Delta w^k)\| > 0$$

et

$$\|\Phi_{\tau_k}(w^k + \Delta w^k)\| \leq \beta \tau_k.$$

La continuité de Φ en w et τ implique qu'il existe un exposant unique l_k tel que $t_k = \rho^{l_k}$ a les propriétés voulues dans l'Etape 2.

□

D'autre part, nous avons la propriété suivante pour la longueur du pas dans l'Etape 3 :

Proposition 1.5 *La longueur du pas \hat{t}_k de l'Etape 3 de l'algorithme 1 est définie de façon unique.*

Introduisons une discussion sur la façon dont nous calculons la longueur de pas $\hat{t}_k > 0$ à l'Etape 3. Cette analyse nous servira un peu plus loin.

Notons que le gradient $\nabla \Psi_{\hat{\tau}_k}(\hat{w}^k)$ vaut

$$\nabla \Psi_{\hat{\tau}_k}(\hat{w}^k) = (\Phi'_{\hat{\tau}_k}(\hat{w}^k))^T \Phi_{\hat{\tau}_k}(\hat{w}^k).$$

De plus, nous obtenons de la proposition 1.1 et du système linéaire (1.21) que

$$\Delta \hat{w}^k = -(\Phi'_{\hat{\tau}_k}(\hat{w}^k))^{-1} \Phi_{\hat{\tau}_k}(\hat{w}^k).$$

La règle d'Armijo (1.22) peut se réécrire sous la forme

$$\Psi_{\hat{\tau}_k}(\hat{w}^k + \hat{t}_k \Delta \hat{w}^k) \leq \Psi_{\hat{\tau}_k}(\hat{w}^k) - 2\hat{t}_k \sigma \Psi_{\hat{\tau}_k}(\hat{w}^k) = (1 - 2\hat{t}_k \sigma) \Psi_{\hat{\tau}_k}(\hat{w}^k). \quad (1.26)$$

En particulier, il suit que l'inégalité

$$\|\Phi_{\hat{\tau}_k}(w^{k+1})\| < \|\Phi_{\hat{\tau}_k}(\hat{w}^k)\| \quad (1.27)$$

est vérifiée pour tout $k \in \mathbb{N}$.

Le prochain résultat montre que tous les itérés w^k déterminés par l'algorithme 1 appartiennent au voisinage

$$\mathcal{N}(\beta) := \{w \mid \|\Phi_{\tau}(w)\| \leq \beta \tau \text{ pour } \tau > 0\}$$

de la trajectoire centrale introduite dans la section précédente.

Proposition 1.6 *Les itérés w^k engendrés par l'algorithme 1 satisfont*

$$\|\Phi_{\tau_k}(w^k)\| \leq \beta\tau_k$$

pour tout $k \in \mathbb{N}$.

Preuve :

Par induction sur k . Pour $k = 0$, le choix de β à l'Etape 0 garantit que

$$\|\Phi_{\tau_0}(w^0)\| \leq \beta\tau_0.$$

Ainsi, supposons que

$$\|\Phi_{\tau_k}(w^k)\| \leq \beta\tau_k.$$

pour $k \geq 0$. Alors, les règles de mise à jour de l'Etape 2 impliquent que

$$\|\Phi_{\hat{\tau}_k}(\hat{w}^k)\| \leq \beta\hat{\tau}_k.$$

Par conséquent, nous obtenons de (1.27)

$$\|\Phi_{\hat{\tau}_k}(w^{k+1})\| < \|\Phi_{\hat{\tau}_k}(\hat{w}^k)\| \leq \beta\hat{\tau}_k.$$

Par continuité (γ_k tend vers 0), la procédure pour calculer $\gamma_k > 0$ de l'Etape 3 est finie et donne un nouveau paramètre de régularisation τ_{k+1} vérifiant

$$\|\Phi_{\tau_{k+1}}(w^{k+1})\| \leq \beta\tau_{k+1}.$$

□

Par les résultats sus-cités, nous déduisons que l'algorithme 1 est bien défini. Enonçons une dernière propriété de l'algorithme 1.

Lemme 1.3 *Supposons que le point de départ $w^0 = (x^0, \lambda^0, s^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ vérifie les équations linéaires*

$$A^T \lambda^0 + s^0 = c \text{ et } Ax^0 = b.$$

Alors, tous les itérés $w^k = (x^k, \lambda^k, s^k)$ produits par l'algorithme 1 satisfont ces équations, i.e., nous avons

$$A^T \lambda^k + s^k = c \text{ et } Ax^k = b$$

pour tout $k \in \mathbb{N}$.

La construction de points $w^0 = (x^0, \lambda^0, s^0)$ remplissant les hypothèses du lemme ci-dessus est aisée car ils ne sont pas supposés avoir des composantes x^0 et s^0 positives. C'est en opposition avec les méthodes des points intérieurs.

1.4 Convergence.

1.4.1 Convergence globale.

Nous allons établir le résultat de convergence globale. Pour y arriver, souvenons-nous de la proposition 1.6. Toute suite $\{w^k\}$ déterminée par l'algorithme 1 vérifie la propriété

$$\|\Phi_{\tau_k}(w^k)\| \leq \beta\tau_k$$

pour tout $k \in \mathbb{N}$. Par conséquent, si τ_k converge vers zéro, nous obtenons

$$\|\Phi_{\tau_k}(w^k)\| \rightarrow 0.$$

Cette relation nous montre que chaque point d'accumulation de la suite $\{w^k\}$ est solution des conditions d'optimalité.

La question qui nous vient à l'esprit est : Sous quelles hypothèses la suite $\{\tau_k\}$ converge-t-elle vers zéro?

Proposition 1.7 *Supposons que la suite $\{w^k\}$ produite par l'algorithme 1 possède au moins un point d'accumulation. Alors, la suite des paramètres de régularisation $\{\tau_k\}$ converge vers zéro.*

Preuve :

Par construction, la suite $\{\tau_k\}$ de paramètres de régularisation est décroissante. Etant également bornée inférieurement, il suit que $\{\tau_k\}$ converge vers un nombre $\bar{\tau} \geq 0$. Si $\bar{\tau} = 0$, c'est fini. Supposons que $\bar{\tau} > 0$. Alors, les règles de mise à jour dans l'Etape 2 de l'algorithme 1 entraînent que

$$\hat{\tau}_k = \tau_k \text{ et } \hat{w}^k = w^k \tag{1.28}$$

pour tout $k \in \mathbb{N}$ suffisamment grand. De plus, par l'Etape 3, nous avons automatiquement

$$\gamma_k \rightarrow 0 \tag{1.29}$$

par la définition de τ_k en fonction de τ_0 . ($\tau_{k+1} = (1 - \gamma_k)\hat{\tau}_k = (1 - \gamma_k)(1 - \gamma_{k-1})\hat{\tau}_{k-1} = \dots$).

Soit w^* un point d'accumulation de la suite $\{w^k\}$ dont l'existence est garantie par notre hypothèse. Soit $\{w^k\}_K$ la sous-suite convergeant vers w^* . Comme $\Phi'_\tau(w^*)$ est non-singulière par la proposition 1.1, des résultats standards de perturbation (cfr [4]), nous trouvons une constante $\widehat{c} > 0$ telle que

$$\|\Phi'_{\tau_k}(w^k)^{-1}\| = \|\Phi'_{\tau_k}(\widehat{w}^k)^{-1}\| \leq \widehat{c} \quad (1.30)$$

pour tout $k \in K$ suffisamment grand.

Considérons deux cas : $\liminf_{k \in K} \widehat{t}_k > 0$ et $\liminf_{k \in K} \widehat{t}_k = 0$ et nous arrivons à une contradiction.

Cas 1 : $\liminf_{k \in K} \widehat{t}_k > 0$.

Alors, il existe une constante positive $\bar{t} > 0$ telle que $\widehat{t}_k \geq \bar{t}$, pour tout $k \in K$. De (1.26), (1.28) et (1.22), nous avons

$$\Psi_{\tau_k}(w^{k+1}) \leq (1 - 2\widehat{t}_k\sigma)\Psi_{\tau_k}(w^k) \leq (1 - 2\bar{t}\sigma)\Psi_{\tau_k}(w^k).$$

La proposition 1.6 nous fournit

$$\begin{aligned} \|\Phi_{\tau_k}(w^{k+1})\| &\leq \sqrt{1 - 2\bar{t}\sigma}\|\Phi_{\tau_k}(w^k)\| \\ &\leq (1 - \bar{c})\|\Phi_{\tau_k}(w^k)\| \\ &\leq \beta(1 - \bar{c})\tau_k \end{aligned} \quad (1.31)$$

pour une constante $\bar{c} \in (0, 1)$. L'inégalité

$$\|\Phi_\tau(w) - \Phi_{\tau'}(w)\| \leq \kappa|\tau - \tau'|$$

est satisfaite pour tout $\tau, \tau' > 0$ et pour tout w où $\kappa > 0$ désigne la constante du lemme 2.5 (cfr [13]). Nous obtenons donc de (1.31)

$$\begin{aligned} \|\Phi_{(1-\gamma)\tau_k}(w^{k+1})\| &\leq \|\Phi_{\tau_k}(w^{k+1})\| + \|\Phi_{(1-\gamma)\tau_k}(w^{k+1}) - \Phi_{\tau_k}(w^{k+1})\| \\ &\leq \beta(1 - \bar{c})\tau_k + \kappa\gamma\tau_k \\ &= (\beta - \bar{c}\beta + \kappa\gamma)\tau_k \end{aligned}$$

pour tout $\gamma \in (0, 1)$. Comme l'inégalité

$$(\beta - \bar{c}\beta + \kappa\gamma)\tau_k \leq \beta(1 - \gamma)\tau_k$$

est vérifiée pour tout $\gamma > 0$ satisfaisant

$$\gamma \leq \bar{c}\beta/(\kappa + \beta),$$

le calcul de γ_k à l'Etape 3 donne

$$\gamma_k \geq \rho \bar{c} \beta / (\kappa + \beta) =: \bar{\gamma}$$

pour tout $k \in K$ suffisamment grand, i.e., ces γ_k sont uniformément bornés par une constante positive $\bar{\gamma} > 0$. C'est en contradiction avec (1.29).

Cas 2 : $\liminf_{k \in K} \hat{t}_k = 0$.

Passant à une sous-suite si nécessaire, nous avons $\lim_{k \in K} \hat{t}_k = 0$. D'où, la condition (1.22) d'Armijo n'est pas satisfaite pour $\alpha_k := \hat{t}_k / \rho$ et pour tout $k \in K$ suffisamment grand, i.e., nous obtenons

$$\Psi_{\hat{\tau}_k}(\hat{w}^k + \alpha_k \Delta \hat{w}^k) > \Psi_{\hat{\tau}_k}(\hat{w}^k) + \alpha_k \sigma \nabla \Psi_{\hat{\tau}_k}(\hat{w}^k)^T \Delta \hat{w}^k$$

ou, de manière équivalente,

$$\frac{\Psi_{\hat{\tau}_k}(\hat{w}^k + \alpha_k \Delta \hat{w}^k) - \Psi_{\hat{\tau}_k}(\hat{w}^k)}{\alpha_k} > \sigma \nabla \Psi_{\hat{\tau}_k}(\hat{w}^k)^T \Delta \hat{w}^k.$$

L'utilisation du Théorème de la Valeur Moyenne et (1.28) nous fournissent

$$\nabla \Psi_{\tau_k}(\xi^k)^T \Delta \hat{w}^k > \sigma \nabla \Psi_{\tau_k}(w^k)^T \Delta \hat{w}^k \quad (1.32)$$

pour un vecteur quelconque ξ^k appartenant au segment de droite joignant w^k et $w^k + \alpha_k \Delta \hat{w}^k$. De (1.30), (1.21), $\{\hat{\tau}_k\}_K = \{\tau_k\}_K \rightarrow \bar{\tau}$ et $\{\hat{w}^k\}_K = \{w^k\}_K \rightarrow w^*$, il suit que la suite $\{\Delta \hat{w}^k\}_K$ converge vers le vecteur Δw^* solution du système linéaire

$$\Phi'_{\bar{\tau}}(w^*) \Delta \hat{w} = -\Phi_{\bar{\tau}}(w^*). \quad (1.33)$$

Prenant la limite $k \rightarrow \infty$ sur K et utilisant $\{\tau_k\}_K \rightarrow \bar{\tau}$ et $\{\alpha_k\}_K \rightarrow 0$, nous arrivons donc à

$$\nabla \Psi_{\bar{\tau}}(w^*)^T \Delta w^* \geq \sigma \nabla \Psi_{\bar{\tau}}(w^*)^T \Delta w^*$$

suite à (1.32) et à la continuité de $\nabla \Psi$ en w et τ . Par conséquent

$$\nabla \Psi_{\bar{\tau}}(w^*)^T \Delta w^* \geq 0.$$

D'autre part, il suit de

$$\nabla \Psi_{\hat{\tau}_k}(\hat{w}^k)^T \Delta \hat{w}^k = -\|\Phi_{\hat{\tau}_k}(\hat{w}^k)\|^2 \leq 0$$

que

$$\nabla \Psi_{\bar{\tau}}(w^*)^T \Delta w^* \leq 0.$$

D'où

$$\nabla \Psi_{\bar{\tau}}(w^*)^T \Delta w^* = 0.$$

Comme $\nabla \Psi_{\bar{\tau}}(w^*) = (\Phi'_{\bar{\tau}}(w^*))^T \Phi_{\bar{\tau}}(w^*)$, de (1.33), nous aboutissons à

$$\Phi_{\bar{\tau}}(w^*) = 0. \quad (1.34)$$

D'un autre côté, il suit de (1.29) que γ_k/ρ ne satisfait pas le test (1.23) pour tout $k \in K$ suffisamment grand, i.e.,

$$\|\Phi_{(1-\gamma_k/\rho)\tau_k}(w^{k+1})\| > \beta \left(1 - \frac{\gamma_k}{\rho}\right) \tau_k.$$

Puisque $\gamma_k/\rho \rightarrow 0$, $\hat{t}_k \rightarrow 0$ et, par conséquent, $w^{k+1} \rightarrow w^*$ pour $k \in K$, nous avons

$$\|\Phi_{\bar{\tau}}(w^*)\| \geq \beta \bar{\tau} > 0,$$

en prenant la limite $k \rightarrow \infty$ sur K . C'est en opposition avec (1.34). \square

L'hypothèse utilisée dans la proposition 1.7, i.e., l'existence d'au moins un point d'accumulation, est plus faible que la condition que la suite entière $\{w^k\}$ reste bornée.

Etablissons le résultat principal de convergence globale de notre algorithme.

Théorème 1.1 *Chaque point d'accumulation de la suite $\{w^k\}$ déterminée par l'algorithme 1 est une solution des conditions d'optimalité (1.3).*

Preuve :

De la proposition 1.6, nous avons

$$\|\Phi_{\tau_k}(w^k)\| \leq \beta \tau_k \quad (1.35)$$

pour tout $k \in \mathbb{N}$. Soit w^* un point d'accumulation de la suite $\{w^k\}$ et soit $\{w^k\}_K$ une sous-suite convergeant vers w^* . Par la proposition 1.7, $\{\tau_k\} \rightarrow 0$. Prenons la limite $k \rightarrow \infty$ pour $k \in K$ et utilisons le fait que Φ est une fonction continue en w et τ , et déduisons de (1.35) que

$$\|\Phi(w^*)\| = \|\Phi_0(w^*)\| \leq 0.$$

En effet,

$$\|\Phi(w^*)\| = \|\Phi_0(w^*)\| = \lim_{k \in K} \|\Phi_{\tau_k}(w^k)\| \leq \lim_{k \in K} \beta \tau_k = 0.$$

D'où,

$$\Phi(w^*) = 0.$$

La définition de Φ entraîne que w^* est une solution des conditions d'optimalité (1.3). □

Le prochain résultat donne une condition suffisante pour l'existence de points d'accumulation.

Théorème 1.2 *Supposons qu'il existe un point strictement admissible pour les conditions d'optimalité (1.3) et que le point de départ $w^0 = (x^0, \lambda^0, s^0)$ vérifie les équations linéaires $Ax^0 = b$ et $A^T\lambda^0 + s^0 = c$. Alors, la suite $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ engendrée par l'algorithme reste bornée.*

Preuve : Utilisant le lemme 1.3 et notre hypothèse sur le point de départ, nous avons

$$A^T\lambda^k + s^k = c \text{ et } Ax^k = b$$

pour tout $k \in \mathbb{N}$. Puisque $\|\Phi(w^k)\| \leq \beta\tau_k$ en vue de la proposition 1.6 et $\tau_k \leq \tau_0$ pour tout $k \in \mathbb{N}$, la thèse découle du corollaire 1.2. □

1.4.2 Convergence locale.

Théorème 1.3 *Nous avons*

$$\tau_{k+1} = O(\tau_k^2)$$

pour tout $k \in \mathbb{N}$ suffisamment grand, i.e., la suite τ_k converge vers zéro au moins Q – quadratiquement.

Toute l'analyse pour parvenir à ce résultat se trouve dans [8]. Remarquons que ce résultat est démontré seulement dans le cas de la fonction minimum. Mais les auteurs pensent que ce théorème est également satisfait dans le cas de la fonction de Fischer-Burmeister mais cela reste une question ouverte.

1.5 Calcul du point de départ et Critère d'arrêt.

Le point initial $w^0 = (x^0, \lambda^0, s^0)$ est construit comme suit :

1. Résoudre $AA^T y = b$ en utilisant la factorisation de Cholesky pour calculer $y^0 \in \mathbb{R}^m$.
2. Poser $x^0 := A^T y^0$.
3. Résoudre $AA^T \lambda = Ac$ en utilisant la factorisation de Cholesky pour calculer $\lambda^0 \in \mathbb{R}^m$.
4. Poser $s^0 := c - A^T \lambda^0$.

Notons que le point de départ est admissible dans le sens où il vérifie les équations linéaires $A^T \lambda^0 + s^0 = c$ et $Ax^0 = b$. Par un tel choix, les composantes de s^0 sont nulles ou petites. Une conséquence est que le résidu initial $\|\Phi(w^0)\|$ est petit. Le point de départ est donc proche de l'ensemble des solutions des conditions d'optimalité. Un autre choix possible pour le point de départ est celui des méthodes des points intérieurs.

1. Résoudre $AA^T y = b$ en utilisant la factorisation de Cholesky pour calculer $y^0 \in \mathbb{R}^m$.
2. Poser $x^0 := A^T y^0$.
3. $\lambda^0 := 0$ et $s^0 := c$.

Dans ce cas, les vecteurs x et s peuvent avoir des composantes négatives.

Le critère d'arrêt.

Nous choisissons $\tau_k < 10^{-4}$ par analogie avec les méthodes des points intérieurs. En effet, ces méthodes suivent la trajectoire centrale (1.18) paramétrisée par μ_k . Le critère d'arrêt est alors $\mu_k < 10^{-8}$. Dans notre approche, nous avons que la condition de complémentarité est paramétrisée par τ^2 . D'où ce choix. Les paramètres restants sont tels que :

$$\tau_0 := \min\{\|\Phi(w^0)\|, 10^5\}, \beta := \|\Phi_{\tau_0}(w^0)\|/\tau_0, \rho = 0.9 \text{ et } \sigma = 10^{-4}.$$

Chapitre 2

Les méthodes régulières Prédicteur-Correcteur pour les programmes linéaires avec une mise à jour plus flexible du paramètre de régularisation.

La façon dont le paramètre est mis à jour a une énorme influence sur le comportement de l'algorithme.

Nous allons décrire la méthode utilisée pour résoudre les conditions d'optimalité (1.3) et montrer que celle-ci est bien définie lors d'une mise à jour plus flexible du paramètre de régularisation. Comme notre approche consiste à exprimer d'une nouvelle manière les conditions d'optimalité, commençons avec une formulation simple de ces conditions. Soit $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction minimum telle que

$$\varphi(a, b) := 2 \min\{a, b\} = a + b - \sqrt{(a - b)^2}$$

et soit $\phi: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ définie par

$$\phi(x, s) := (\varphi(x_1, s_1), \dots, \varphi(x_n, s_n))^T.$$

Comme φ possède la propriété

$$\varphi(a, b) = 0 \text{ ssi } a \geq 0, b \geq 0, ab = 0,$$

nous en déduisons que ϕ peut être utilisée pour obtenir une caractérisation des conditions de complémentarité, i.e.,

$$x_i \geq 0, s_i \geq 0, x_i s_i = 0 \text{ ssi } \phi(x, s) = 0.$$

D'où, un vecteur $w^* := (x^*, \lambda^*, s^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ est solution des conditions d'optimalité si et seulement s'il satisfait le système non linéaire d'équations $\Phi(w) = 0$ où $\Phi : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ est donnée par

$$\Phi(w) := \Phi(x, \lambda, s) := \begin{pmatrix} A^T \lambda + s - c \\ Ax - b \\ \phi(x, s) \end{pmatrix}.$$

Le principal inconvénient est que cette fonction n'est pas différentiable partout. Pour résoudre ce problème, il est judicieux d'approximer la fonction minimum φ par une fonction continûment différentiable à l'aide d'un paramètre de régularisation τ . En particulier, la fonction

$$\varphi_\tau(a, b) := a + b - \sqrt{(a - b)^2 + 4\tau^2}.$$

Grâce à cette fonction, nous définissons les fonctions

$$\phi_\tau(x, s) := (\varphi_\tau(x_1, s_1), \dots, \varphi_\tau(x_n, s_n))^T$$

et

$$\Phi_\tau(w) := \Phi_\tau(x, \lambda, s) := \begin{pmatrix} A^T \lambda + s - c \\ Ax - b \\ \phi_\tau(x, s) \end{pmatrix}.$$

Donc Φ_τ est une approximation régulière de Φ pour chaque $\tau > 0$. Grâce à la référence [12], nous voyons que le vecteur $w_\tau = (x_\tau, \lambda_\tau, s_\tau)$ est solution du système non linéaire d'équations

$$\Phi_\tau(w) = 0 \tag{2.1}$$

ssi ce vecteur résout les conditions de la trajectoire centrale (1.18).

A partir de maintenant, considérons τ comme une variable indépendante et introduisons les notations suivantes : $\varphi(a, b, \tau) := \varphi_\tau(a, b)$ et $\phi(x, s, \tau) := \phi_\tau(x, s)$. Le système non linéaire (2.1) a $n+m+n$ équations et $\Phi(x, \lambda, s, \tau) := \Phi_\tau(x, \lambda, s)$ possède $n + m + n + 1$ variables. Si nous ajoutons une ligne à ce système, nous pouvons définir la fonction :

$$\Theta(x, \lambda, s, \tau) := \begin{pmatrix} A^T \lambda + s - c \\ Ax - b \\ \phi(x, s, \tau) \\ \tau \end{pmatrix}.$$

Nous aurons besoin de la généralisation suivante de la fonction Θ

$$\Theta_{\sigma, \psi}(x, \lambda, s, \tau) := \begin{pmatrix} A^T + s - c \\ Ax - b \\ \phi(x, s, \tau) \\ \sigma\psi(\tau) \end{pmatrix},$$

où $\sigma \in (0, 1]$ est un paramètre de centrage et $\psi: [0, \infty) \rightarrow \mathbb{R}$ est une fonction possédant les propriétés suivantes :

- (1) ψ est continûment différentiable avec $\psi(0) = 0$.
- (2) $\psi'(\tau) > 0$ pour tout $\tau \geq 0$.
- (3) $\psi(\tau) \leq \psi'(\tau)\tau$ pour tout $\tau \geq 0$.
- (4) Pour chaque $\tau_0 > 0$, il existe une constante $\gamma > 0$ (pouvant dépendre de τ_0) telle que $\psi(\tau) \geq \gamma\psi'(\tau)\tau$ pour tout $\tau \in [0, \tau_0]$.

Voici quelques fonctions satisfaisant ces propriétés avec la valeur de γ correspondante :

1. $\psi(\tau) := \tau$ où $\gamma = 1$,
2. $\psi(\tau) := (1 + \tau)^2 - 1$ où $\gamma = 1/2$,
3. $\psi(\tau) := \exp(\tau) - 1$ où $\gamma = (1 - \exp(-\tau_0))/\tau_0$.

Faisons quelques commentaires sur les propriétés de la fonction ψ :

La propriété (1) est indispensable puisque nous voulons appliquer une méthode de type Newton au système d'équations $\Theta_{\sigma, \psi}(x, \lambda, s, \tau) = 0$. D'où ψ doit être suffisamment différentiable.

La propriété (2) implique que ψ est strictement croissante. Combiné avec $\psi(0) = 0$, cela signifie que le système non linéaire d'équations $\Theta_{\sigma, \psi}(x, \lambda, s, \tau) = 0$ est équivalent aux conditions d'optimalité.

La propriété (3) sera utilisée pour montrer que l'algorithme 2 est bien défini. De plus, les propriétés (3) et (4) garantiront la décroissance monotone de la suite $\{\tau_k\}$ et sa convergence vers zéro.

La méthode décrite ici est un algorithme Prédicteur-Correcteur. Le pas Prédicteur est responsable d'un taux de convergence locale rapide tandis que le pas Correcteur garantit la convergence globale. Plus précisément, le pas Prédicteur est une itération de Newton appliquée au système $\Theta(x, \lambda, s, \tau) = 0$ suivie par une mise à jour adéquate de τ . Elle a pour but de réduire τ au maximum. Le pas Correcteur applique une itération Newton du système $\Theta_{1, \psi}(x, \lambda, s, \tau) = 0$. Le membre de droite habituel $\Theta_{1, \psi}(x, \lambda, s, \tau)$ est remplacé par $\Theta_{\sigma, \psi}(x, \lambda, s, \tau)$ où $\sigma \in (0, 1)$ est un paramètre de centrage. Ce pas de Newton est suivi par une recherche linéaire de type Armijo.

Le calcul des itérés est tel qu'ils appartiennent au voisinage de la trajectoire centrale

$$\mathcal{N}(\beta) := \{(x, \lambda, s, \tau) | A^T \lambda + s = c, Ax = b, \|\phi(x, s, \tau)\| \leq \beta \tau\}. \quad (2.2)$$

Algorithme 2.

Etape 0: Initialisation.

Choisir $w^0 := (x^0, \lambda^0, s^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ et $\tau_0 > 0$ tels que $A^T \lambda^0 + s^0 = c$, $Ax^0 = b$ et $\phi(x^0, s^0, \tau_0) \leq 0$.

Choisir $\beta \geq \|\phi(x^0, s^0, \tau_0)\|/\tau_0$, $\rho \in (0, 1)$, $0 < \hat{\sigma}_{\min} < \hat{\sigma}_{\max} < 1$, $\epsilon \geq 0$ et poser $k := 0$.

Etape 1: Critère d'arrêt.

Si $\|\phi(x^k, s^k, 0)\| \leq \epsilon$: Stop.

Etape 2: Pas Prédicteur.

Calculer une solution $(\Delta w^k, \Delta \tau_k) = (\Delta x^k, \Delta \lambda^k, \Delta s^k, \Delta \tau_k) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}$ du système linéaire

$$\Theta'(w^k, \tau_k) \begin{pmatrix} \Delta w \\ \Delta \tau \end{pmatrix} = -\Theta(w^k, \tau_k). \quad (2.3)$$

Si $\|\Phi(x^k + \Delta x^k, s^k + \Delta s^k, 0)\| = 0$: STOP.

Sinon si

$$\|\Phi(x^k + \Delta x^k, s^k + \Delta s^k, \tau_k)\| > \beta \tau_k$$

alors poser

$$\hat{w}^k := w^k, \hat{\tau}_k := \tau_k \text{ et } \eta_k := 1,$$

sinon calculer $\eta_k = \rho^{l_k}$ où l_k est un entier positif ou nul tel que

$$\|\Phi(x^k + \Delta x^k, s^k + \Delta s^k, \rho^j \tau_k)\| \leq \beta \rho^j \tau_k \quad \forall j = 0, 1, 2, \dots, l_k \text{ et}$$

$$\|\Phi(x^k + \Delta x^k, s^k + \Delta s^k, \rho^{l_k+1} \tau_k)\| > \beta \rho^{l_k+1} \tau_k,$$

et poser

$$\hat{\tau}_k := \eta_k \tau_k \text{ et } \hat{w}^k := \begin{cases} w^k & \text{si } l_k = 0 \\ w^k + \Delta w^k & \text{sinon.} \end{cases}$$

Etape 3 : Pas Correcteur.

Choisir $\hat{\sigma}_k \in [\hat{\sigma}_{\min}, \hat{\sigma}_{\max}]$ et calculer une solution $(\Delta \hat{w}^k, \Delta \hat{\tau}_k)$
 $= (\Delta \hat{x}^k, \Delta \hat{\lambda}^k, \Delta \hat{s}^k, \Delta \hat{\tau}_k) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}$ du système linéaire

$$\Theta'_{1, \psi}(\hat{w}^k, \hat{\tau}_k) \begin{pmatrix} \Delta \hat{w} \\ \Delta \hat{\tau} \end{pmatrix} = -\Theta_{\hat{\sigma}_k, \psi}(\hat{w}^k, \hat{\tau}_k). \quad (2.4)$$

Soit $\hat{t}_k = \max\{\rho^l \mid l = 0, 1, 2, \dots\}$ tel que

$$\|\Phi(\hat{x}^k + \hat{t}_k \Delta \hat{x}^k, \hat{s}^k + \hat{t}_k \Delta \hat{s}^k, \hat{\tau}_k + \hat{t}_k \Delta \hat{\tau}_k)\| \leq \beta(\hat{\tau}_k + \hat{t}_k \Delta \hat{\tau}_k) \quad (2.5)$$

et poser

$$w^{k+1} := \hat{w}^k + \hat{t}_k \Delta \hat{w}^k \text{ et } \tau_{k+1} := \hat{\tau}_k + \hat{t}_k \Delta \hat{\tau}_k.$$

Etape 4 : Mise à jour.

Poser $k = k + 1$ et aller à l'Etape 1.

Nous supposons dans la suite que le paramètre ϵ de l'Etape 1 vaut zéro. Si l'algorithme 2 engendre une suite infinie, les critères d'arrêt des Etapes 2 et 3 ne seront pas satisfaits.

Montrons que l'algorithme 2 est bien défini.

Lemme 2.1 *Les énoncés suivants sont vérifiés pour tout $k \in \mathbb{N}$:*

1. *Les systèmes linéaires (2.3) et (2.4) ont une solution unique.*
2. *Il existe un seul η_k satisfaisant les conditions de l'Etape 2.*
3. *La taille du pas \hat{t}_k de l'Etape 3 est définie de manière unique.*

Par conséquent, l'algorithme 2 est bien défini.

Preuve :

1. La preuve est une conséquence de la proposition 3.1 de la référence [6], si nous tenons compte de la structure des matrices Jacobiennes $\Theta'(w, \tau)$ et $\Theta'_{\sigma, \psi}(w, \tau)$ et de la propriété (2) de ψ .

2. La preuve est analogue à celle faite dans le premier chapitre. (voir la proposition 1.4).

3. Supposons qu'il existe une itération k où

$$\|\Phi(\hat{x}^k + \rho^l \Delta \hat{x}^k, \hat{s}^k + \rho^l \Delta \hat{s}^k, \hat{\tau}_k + \rho^l \Delta \hat{\tau}_k)\| > \beta(\hat{\tau}_k + \rho^l \Delta \hat{\tau}_k),$$

pour tout $l \in \mathbb{N}$. Comme

$$\|\Phi(\hat{x}^k, \hat{s}^k, \hat{\tau}_k)\| \leq \beta \hat{\tau}_k,$$

nous obtenons de la propriété (3) de ψ

$$\begin{aligned} \beta(\hat{\tau}_k + \rho^l \Delta \hat{\tau}_k) &= \beta \left(\hat{\tau}_k - \rho^l \hat{\sigma}_k \frac{\psi(\hat{\tau}_k)}{\psi'(\hat{\tau}_k)} \right) \\ &\geq \beta(\hat{\tau}_k - \rho^l \hat{\sigma}_k \hat{\tau}_k) \\ &\geq \beta(1 - \hat{\sigma}_{\max} \rho^l) \hat{\tau}_k \\ &\geq (1 - \hat{\sigma}_{\max} \rho^l) \|\phi(\hat{x}^k, \hat{s}^k, \hat{\tau}_k)\|. \end{aligned}$$

Prenant cette inégalité en compte, nous avons

$$\|\Phi(\hat{x}^k + \rho^l \Delta \hat{x}^k, \hat{s}^k + \rho^l \Delta \hat{s}^k, \hat{\tau}_k + \rho^l \Delta \hat{\tau}_k)\| > (1 - \hat{\sigma}_{\max} \rho^l) \|\phi(\hat{x}^k, \hat{s}^k, \hat{\tau}_k)\|.$$

Or, par Armijo, nous obtenons

$$\|\Phi(\hat{x}^k + \rho^l \Delta \hat{x}^k, \hat{s}^k + \rho^l \Delta \hat{s}^k, \hat{\tau}_k + \rho^l \Delta \hat{\tau}_k)\| \leq (1 - \hat{\sigma}_{\max} \rho^l) \|\phi(\hat{x}^k, \hat{s}^k, \hat{\tau}_k)\|.$$

C'est en opposition avec l'hypothèse. D'où, la longueur du pas à l'Étape 3 est définie de façon unique.

□

Etablissons quelques propriétés de l'algorithme 2.

Lemme 2.2 *Les suites $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ et $\{\tau_k\}$ déterminées par l'algorithme 2 vérifient les propriétés suivantes :*

1. $A^T \lambda^k + s^k = c$ et $Ax^k = b$ pour tout $k \in \mathbb{N}$.
2. $\tau_k \leq \tau_0(1 - \gamma \hat{\sigma}_0 \hat{t}_0) \eta_0 \cdots (1 - \gamma \hat{\sigma}_{k+1} \hat{t}_{k+1}) \eta_{k+1}$ pour tout $k \in \mathbb{N}$ et $\gamma > 0$ est la constante de la propriété (4) de la fonction ψ .
3. $\|\phi(x^k, s^k, \tau_k)\| \leq \beta \tau_k$ pour tout $k \in \mathbb{N}$.

Preuve :

1. Pour $k = 0$, la propriété découle du choix du point de départ. Pour $k >$

0, c'est une conséquence de la résolution par la méthode de Newton des systèmes linéaires.

2. Nous obtenons

$$\Delta\tau_k = -\hat{\sigma}_k\psi(\hat{\tau}_k)/\psi'(\hat{\tau}_k) \quad (2.6)$$

de la quatrième équation du système (2.4) de l'algorithme 2. Comme $\tau_k, \hat{\tau}_k \in [0, \tau_0]$ pour tout $k \in \mathbb{N}$, il suit de la propriété (4) de la fonction ψ et des règles de mise à jour dans les Etapes 2 et 3 que

$$\begin{aligned} \tau_{k+1} &= \hat{\tau}_k + \hat{t}_k \Delta\hat{\tau}_k \\ &= \hat{\tau}_k - \hat{t}_k \hat{\sigma}_k \psi(\hat{\tau}_k) / \psi'(\hat{\tau}_k) \\ &\leq \hat{\tau}_k - \gamma \hat{t}_k \hat{\sigma}_k \hat{\tau}_k \\ &= (1 - \gamma \hat{t}_k \hat{\sigma}_k) \eta_k \tau_k. \end{aligned}$$

La thèse découle d'une simple induction sur k .

3. C'est une conséquence directe des règles de mise à jour de l'algorithme 2. \square

2.1 Convergence.

Nous analyserons les propriétés de convergence locale et globale de l'algorithme 2 sus-cité. Nous allons montrer que la suite des itérés (x^k, λ^k, s^k) est bornée sous l'hypothèse d'admissibilité stricte. Ce résultat sera prouvé en montrant que les itérés vérifient l'inégalité

$$\phi(x^k, s^k, \tau_k) \leq 0$$

pour tout $k \in \mathbb{N}$.

Lemme 2.3 *Les suites $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$, $\{\tau_k\}$, $\{\hat{w}^k\}$ et $\{\hat{\tau}_k\}$ engendrées par l'algorithme 2 ont les propriétés*

1. $\phi(\hat{x}^k, \hat{s}^k, \hat{\tau}_k) \leq 0$ pour tout $k \in \mathbb{N}$.
2. $\phi(x^k, s^k, \tau_k) \leq 0$ pour tout $k \in \mathbb{N}$.

Preuve :

Considérons d'abord l'énoncé 1. Supposons k fixé et que $\widehat{w}^k = w^k + \Delta w^k$ dans l'Etape 2 de l'algorithme 2. Comme chaque composante de la fonction ϕ est concave, nous obtenons

$$\begin{aligned}
& \phi(\widehat{x}^k, \widehat{s}^k, \widehat{\tau}_k) \\
&= \phi(x^k + \Delta \widehat{x}^k, s^k + \Delta \widehat{s}^k, \eta_k \tau_k) \\
&= \phi(x^k + \Delta \widehat{x}^k, s^k + \Delta \widehat{s}^k, \tau_k + (\eta_k - 1)\tau_k) \\
&\leq \phi(x^k, s^k, \tau_k) + \phi'(x^k, s^k, \tau_k) \begin{pmatrix} \Delta x^k \\ \Delta s^k \\ (\eta_k - 1)\tau_k \end{pmatrix} \\
&= \phi(x^k, s^k, \tau_k) + \phi'(x^k, s^k, \tau_k) \begin{pmatrix} \Delta x^k \\ \Delta s^k \\ \Delta \tau_k \end{pmatrix} \\
&\quad + \phi'(x^k, s^k, \tau_k) \begin{pmatrix} 0 \\ 0 \\ (\eta_k - 1)\tau_k - \Delta \tau_k \end{pmatrix}.
\end{aligned} \tag{2.7}$$

Par la troisième équation du système linéaire (2.3) de l'algorithme 2, nous avons

$$\phi'(x^k, s^k, \tau_k) \begin{pmatrix} \Delta x^k \\ \Delta s^k \\ \Delta \tau_k \end{pmatrix} = -\phi(x^k, s^k, \tau_k).$$

D'où, (2.7) devient

$$\phi(\widehat{x}^k, \widehat{s}^k, \widehat{\tau}_k) \leq \phi'(x^k, s^k, \tau_k) \begin{pmatrix} 0 \\ 0 \\ (\eta_k - 1)\tau_k - \Delta \tau_k \end{pmatrix}. \tag{2.8}$$

Nous prétendons que le membre de droite de (2.8) est négatif ou nul. Pour le prouver, remarquons que

$$\phi'(x^k, s^k, \tau_k) \begin{pmatrix} 0 \\ 0 \\ (\eta_k - 1)\tau_k - \Delta \tau_k \end{pmatrix} = ((\eta_k - 1)\tau_k - \Delta \tau_k) d_\tau^k$$

où

$$\begin{aligned} d_\tau^k &:= \left(\frac{\partial \varphi}{\partial \tau}(x_1^k, s_1^k, \tau_k), \dots, \frac{\partial \varphi}{\partial \tau}(x_n^k, s_n^k, \tau_k) \right)^T \\ &= \left(\frac{-4\tau_k}{\sqrt{(x_1^k - s_1^k)^2 + 4\tau_k^2}}, \dots, \frac{-4\tau_k}{\sqrt{(x_n^k - s_n^k)^2 + 4\tau_k^2}} \right)^T \leq 0. \end{aligned}$$

D'où, il reste à montrer que

$$(\eta_k - 1)\tau_k - \Delta\tau_k \geq 0.$$

C'est évident puisque la dernière ligne du système linéaire (2.3) de l'algorithme 2 donne $\Delta\tau_k = -\tau_k$.

Considérons ensuite l'énoncé 2. Supposant toujours que $k \in \mathbb{N}$ est fixé et sachant que chaque composante de la fonction ϕ est concave, nous avons

$$\begin{aligned} \phi(x^{k+1}, s^{k+1}, \tau_{k+1}) &= \phi(\hat{x}^k + \hat{t}_k \Delta \hat{x}^k, \hat{s}^k + \hat{t}_k \Delta \hat{s}^k, \hat{\tau}_k + \hat{t}_k \Delta \hat{\tau}_k) \\ &\leq \phi(\hat{x}^k, \hat{s}^k, \hat{\tau}_k) + \hat{t}_k \phi'(\hat{x}^k, \hat{s}^k, \hat{\tau}_k) \begin{pmatrix} \Delta \hat{x}^k \\ \Delta \hat{s}^k \\ \Delta \hat{\tau}_k \end{pmatrix} \quad (2.9) \\ &= \phi(\hat{x}^k, \hat{s}^k, \hat{\tau}_k) - \hat{t}_k \phi(\hat{x}^k, \hat{s}^k, \hat{\tau}_k) \\ &= (1 - \hat{t}_k) \phi(\hat{x}^k, \hat{s}^k, \hat{\tau}_k) \leq 0. \end{aligned}$$

Nous allons maintenant vérifier les énoncés 1 et 2 par induction sur k . Si $k = 0$, nous avons $\phi(x^0, s^0, \tau_0) \leq 0$ par le choix du point de départ $w^0 = (x^0, \lambda^0, s^0)$ et du paramètre de régularisation $\tau_0 > 0$ dans l'Etape 0 de l'algorithme 2. Par conséquent, si nous posons $\hat{w}^0 = w^0$ dans l'Etape 2, nous avons $\hat{\tau}_0 = \tau_0$ et $\phi(\hat{x}^0, \hat{s}^0, \hat{\tau}_0) \leq 0$. D'autre part, si nous posons $\hat{w}^0 = w^0 + \Delta w^0$ dans l'Etape 2, l'argument utilisé au début de la preuve montre que l'inégalité $\phi(\hat{x}^0, \hat{s}^0, \hat{\tau}_0) \leq 0$ est aussi satisfaite dans ce cas.

Supposons que nous avons $\phi(x^k, s^k, \tau_k) \leq 0$ et $\phi(\hat{x}^k, \hat{s}^k, \hat{\tau}_k) \leq 0$ pour un certain indice $k \in \mathbb{N}$. Alors, (2.9) implique $\phi(x^{k+1}, s^{k+1}, \tau_{k+1}) \leq 0$. Par conséquent, si nous avons $\hat{w}^{k+1} = w^{k+1}$ à l'Etape 2 de l'algorithme 2, nous obtenons $\phi(\hat{x}^{k+1}, \hat{s}^{k+1}, \hat{\tau}_{k+1}) \leq 0$. Par contre, si $\hat{w}^{k+1} = w^{k+1} + \Delta w^{k+1}$ dans l'Etape 2, l'argument utilisé dans le début de la preuve implique que l'inégalité est satisfaite.

□

Nous allons maintenant montrer que la suite $\{w^k\}$ engendrée par l'algorithme 2 est bornée lorsque les conditions d'optimalité (1.3) admettent un point strictement admissible, i.e., un vecteur $\hat{w} = (\hat{x}, \hat{\lambda}, \hat{s})$ satisfaisant $A^T \hat{\lambda} + \hat{s} = c$, $A\hat{x} = b$ et $\hat{x} > 0$, $\hat{s} > 0$.

Proposition 2.1 *Supposons qu'il existe un point strictement admissible pour les conditions d'optimalité. Alors, la suite $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ engendrée par l'algorithme 2 est bornée.*

Preuve :

Supposons que la suite $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ engendrée par l'algorithme 2 n'est pas bornée. Comme la suite $\{\tau_k\}$ est décroissante par le point 2 du lemme 2.2, il suit du point 3 de ce même lemme que

$$\|\phi(x^k, s^k, \tau_k)\| \leq \beta \tau_k \leq \beta \tau_0 \quad \forall k \in \mathbb{N}. \quad (2.10)$$

Par la définition de la fonction régulière minimum, il ne peut exister d'indice $i \in \{1, \dots, n\}$ tel que $x_i^k \rightarrow -\infty$ ou $s_i^k \rightarrow -\infty$ pour une sous-suite. Sinon, nous aurions $\phi(x_i^k, s_i^k, \tau_k) \rightarrow -\infty$ donc $\|\phi(x^k, s^k, \tau_k)\| \rightarrow +\infty$ ce qui contredit (2.10). D'où, les composantes de ces deux suites sont bornées inférieurement, i.e.,

$$x_i^k \geq \gamma \text{ et } s_i^k \geq \gamma \quad \forall i = 1, \dots, n, \forall k \in \mathbb{N}. \quad (2.11)$$

où γ est une constante réelle.

D'autre part, $\{w^k\}$ est non bornée par hypothèse. Il existe donc au moins un indice $i \in \{1, \dots, n\}$ tel que $x_i^k \rightarrow +\infty$ ou $s_i^k \rightarrow +\infty$ pour une sous-suite. Dans le cas contraire, les deux suites $\{x^k\}$ et $\{s^k\}$ seraient bornées. Cela impliquerait que la suite $\{\lambda^k\}$ le serait aussi par le point 1 du lemme 2.2 et par le fait que A soit de rang plein. Donc la suite $\{w^k\}$ serait bornée ce qui est contraire à l'hypothèse.

Soit $\hat{w} = (\hat{x}, \hat{\lambda}, \hat{s}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ un point strictement admissible pour les conditions d'optimalité dont l'existence est garantie par notre hypothèse. Alors,

$$A^T \hat{\lambda} + \hat{s} = c \text{ et } A\hat{x} = b.$$

Par le point 1 du lemme 2.2, nous avons

$$A^T \lambda^k + s^k = c \text{ et } Ax^k = b \quad \forall k \in \mathbb{N}$$

d'où, nous obtenons par soustraction des deux équations

$$A^T(\widehat{\lambda} - \lambda^k) + (\widehat{s} - s^k) = 0 \text{ et } A(\widehat{x} - x^k) = 0. \quad (2.12)$$

Prémultipliant la première équation de (2.12) par $(\widehat{x} - x^k)^T$ et en tenant compte de la deuxième équation, nous avons

$$(\widehat{x} - x^k)^T (\widehat{s} - s^k) = 0.$$

La réorganisation des termes donne

$$\widehat{s}^T x^k + \widehat{x}^T s^k = (x^k)^T s^k + \widehat{x}^T \widehat{s} \quad \forall k \in \mathbb{N}. \quad (2.13)$$

Par (2.11) et le fait que $\widehat{x} > 0$ et $\widehat{s} > 0$ car \widehat{w} est strictement admissible, nous avons

$$(x^k)^T s^k \rightarrow +\infty$$

suite à (2.13) et au fait que $x_i^k \rightarrow +\infty$ ou $s_i^k \rightarrow +\infty$ pour au moins un indice $i \in \{1, \dots, n\}$ d'une sous-suite. D'où, il existe une composante $j \in \{1, \dots, n\}$ telle que

$$x_j^k s_j^k \rightarrow +\infty \quad (2.14)$$

pour une sous-suite adéquate.

Par le point 2 du lemme 2.3, nous savons que

$$\phi(x^k, s^k, \tau_k) \leq 0 \quad \forall k \in \mathbb{N}.$$

Considérant la j ème composante et la définition de ϕ , nous avons

$$x_j^k + s_j^k \leq \sqrt{(x_j^k - s_j^k)^2 + 4\tau_k^2} \quad (2.15)$$

pour tout $k \in \mathbb{N}$. Par (2.14) et (2.11), nous avons nécessairement $x_j^k > 0$ et $s_j^k > 0$ pour tous ces k appartenant à la sous-suite pour laquelle (2.14) est vérifiée. Si nous élevons les deux membres de (2.15) au carré, nous obtenons après simplification

$$4x_j^k s_j^k \leq 4\tau_k^2.$$

Cependant, le membre de droite de cette équation est bornée par $4\tau_0^2$. C'est en opposition avec (2.14).

□

Prouvons le résultat de convergence globale de l'algorithme 2.

Théorème 2.1 *Supposons que la suite $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ déterminée par l'algorithme 2 admet au moins un point d'accumulation. Alors, la suite $\{\tau_k\}$ converge vers zéro.*

Preuve :

La suite $\{\tau_k\}$ de paramètres de régularisation est décroissante. Etant également bornée inférieurement, il suit que $\{\tau_k\}$ converge vers un nombre $\tau_* \geq 0$. Si $\tau_* = 0$, c'est fini. Supposons que $\tau_* > 0$. Alors, les règles de mise à jour de l'Etape 2 de l'algorithme 2 impliquent

$$\hat{\tau}_k = \tau_k, \eta_k = 1 \text{ et } \hat{w}^k = w^k \quad (2.16)$$

pour tout $k \in \mathbb{N}$ suffisamment grand. Passant à une sous-suite si nécessaire, nous supposons, sans perdre de généralité, que (2.16) est satisfaite pour tout $k \in \mathbb{N}$. Alors, le point 2 du lemme 2.2 et $\hat{\sigma}_k \geq \hat{\sigma}_{\min}$ nous donnent

$$\tau_k \leq \tau_0 \prod_{j=0}^{k-1} (1 - \gamma \hat{\sigma}_j \hat{t}_j) \leq \tau_0 \prod_{j=0}^{k-1} (1 - \gamma \hat{\sigma}_{\min} \hat{t}_j). \quad (2.17)$$

Comme $\tau_k \rightarrow \tau_* > 0$ par hypothèse, il suit de (2.17) que $\lim_{k \rightarrow \infty} \hat{t}_k = 0$. La longueur du pas $\hat{\alpha}_k := \hat{t}_k / \rho$ ne satisfait pas le critère de recherche linéaire (2.5) de l'algorithme 2 pour $k \in \mathbb{N}$ suffisamment grand. D'où,

$$\|\Phi(\hat{x}^k + \hat{\alpha}_k \Delta \hat{x}^k, \hat{s}^k + \hat{\alpha}_k \Delta \hat{s}^k, \hat{\tau}_k + \hat{\alpha}_k \Delta \hat{\tau}_k)\| > \beta(\hat{\tau}_k + \hat{\alpha}_k \Delta \hat{\tau}_k) \quad (2.18)$$

pour tous ces $k \in \mathbb{N}$.

Soit $w^* = (x^*, \lambda^*, s^*)$ un point d'accumulation de la suite $\{w^k\}$ et soit $\{w^k\}_K$ une sous-suite convergeant vers w^* . Comme $\hat{\sigma}_k \in [\hat{\sigma}_{\min}, \hat{\sigma}_{\max}]$ pour tout $k \in \mathbb{N}$, nous pouvons supposer, sans perdre de généralité, que la sous-suite $\{\hat{\sigma}_k\}_K$ converge vers un nombre $\hat{\sigma}_* \in [\hat{\sigma}_{\min}, \hat{\sigma}_{\max}]$. De plus, comme $\tau_* > 0$, il suit de (2.16) et du point 1 du lemme 2.1 que la sous-suite correspondante $\{(\Delta \hat{w}^k, \Delta \hat{\tau}_k)\}_K$ converge vers un vecteur $(\Delta \hat{w}^*, \Delta \hat{\tau}_*) = (\Delta \hat{x}^*, \Delta \hat{\lambda}^*, \Delta \hat{s}^*, \Delta \hat{\tau}_*)$, où $(\Delta \hat{w}^*, \Delta \hat{\tau}_*)$ est la solution unique du système linéaire

$$\Theta'_{1, \psi}(w^*, \tau_*) \begin{pmatrix} \Delta \hat{w} \\ \Delta \hat{\tau} \end{pmatrix} = -\Theta_{\hat{\sigma}_*, \psi}(w^*, \tau_*) \quad (2.19)$$

par analogie au système (2.4) de l'algorithme 2. En utilisant $\{\hat{\alpha}_k\}_K \rightarrow 0$ et prenant la limite $k \rightarrow \infty$ sur le sous-ensemble K , nous obtenons de (2.16) et (2.18)

$$\|\phi(x^*, s^*, \tau_*)\| \geq \beta_{\tau_*} > 0. \quad (2.20)$$

D'autre part, nous obtenons de (2.18), (2.6), la propriété 3 de la fonction ψ , (2.16), le point 3 du lemme 2.2 et $\hat{\sigma}_k \leq \hat{\sigma}_{\max}$ que

$$\begin{aligned} & \|\Phi(\hat{x}^k + \hat{\alpha}_k \Delta \hat{x}^k, \hat{s}^k + \hat{\alpha}_k \Delta \hat{s}^k, \hat{\tau}_k + \hat{\alpha}_k \Delta \hat{\tau}_k)\| \\ & > \beta(\hat{\tau}_k + \hat{\alpha}_k \Delta \hat{\tau}_k) \\ & = \beta(\hat{\tau}_k - \hat{\alpha}_k \hat{\sigma}_k \psi(\hat{\tau}_k) / \psi'(\hat{\tau}_k)) \\ & \geq \beta(\hat{\tau}_k - \hat{\alpha}_k \hat{\sigma}_k \hat{\tau}_k) \\ & = (1 - \hat{\alpha}_k \hat{\sigma}_k) \beta \tau_k \\ & \geq (1 - \hat{\alpha}_k \hat{\sigma}_k) \|\phi(x^k, s^k, \tau_k)\| \\ & \geq (1 - \hat{\alpha}_k \hat{\sigma}_{\max}) \|\phi(x^k, s^k, \tau_k)\| \end{aligned}$$

pour tout $k \in \mathbb{N}$ suffisamment grand. De (2.16), nous obtenons

$$\frac{\|\Phi(\hat{x}^k + \hat{\alpha}_k \Delta \hat{x}^k, \hat{s}^k + \hat{\alpha}_k \Delta \hat{s}^k, \hat{\tau}_k + \hat{\alpha}_k \Delta \hat{\tau}_k)\| - \|\phi(x^k, s^k, \tau_k)\|}{\hat{\alpha}_k} \geq -\hat{\sigma}_{\max} \|\phi(x^k, s^k, \tau_k)\|.$$

Comme $\|\phi(\cdot, \cdot, \cdot)\|$ est continûment différentiable en (x^*, s^*, τ^*) par (2.20), prendre la limite $k \rightarrow \infty$ pour $k \in K$ nous donne

$$\frac{(\phi(x^*, s^*, \tau_*))^T}{\|\phi(x^*, s^*, \tau_*)\|} \phi'(x^*, s^*, \tau_*) \begin{pmatrix} \Delta \hat{x}^* \\ \Delta \hat{s}^* \\ \Delta \hat{\tau}_* \end{pmatrix} \geq -\hat{\sigma}_{\max} \|\phi(x^*, s^*, \tau_*)\|,$$

où $(\Delta \hat{w}^*, \Delta \hat{\tau}_*) = (\Delta \hat{x}^*, \Delta \hat{\lambda}^*, \Delta \hat{s}^*, \Delta \hat{\tau}_*)$ est la solution du système linéaire (2.19). Par (2.19), nous avons

$$-\|\phi(x^*, s^*, \tau_*)\| \geq -\hat{\sigma}_{\max} \|\phi(x^*, s^*, \tau_*)\|.$$

Comme $\hat{\sigma}_{\max} \in (0, 1)$, cela implique que $\|\phi(x^*, s^*, \tau_*)\| = 0$ ce qui contredit (2.20). D'où, nous ne pouvons avoir $\tau_* > 0$.

□

Par la proposition 2.1, l'existence présumée d'un point d'accumulation dans le théorème 2.1 est automatiquement satisfaite s'il a un point strictement admissible pour les conditions d'optimalité. Une conséquence immédiate du théorème 2.1 est le résultat suivant :

Corollaire 2.1 *Chaque point d'accumulation de la suite $\{w^k\}$ déterminée par l'algorithme 2 est une solution des conditions d'optimalité.*

Preuve :

Soit $w^* = (x^*, \lambda^*, s^*)$ un point d'accumulation de la suite $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ et soit $\{w^k\}_K$ une sous-suite convergeant vers w^* . Alors, nous avons $\tau_k \rightarrow 0$ en vue du théorème 2.1. D'où, le point 3 du lemme 2.2 implique

$$\|\phi(x^*, s^*, 0)\| = \lim_{k \in K} \|\phi(x^k, s^k, \tau_k)\| \leq \beta \lim_{k \in K} \tau_k = 0,$$

i.e., nous avons $x^* \geq 0$, $s^* \geq 0$ et $x_i^* s_i^* = 0$ pour $i = 1, \dots, n$ par la définition de ϕ . Puisque le point 3 du lemme 2.2 montre que $A^T \lambda^* + s^* = c$ et $Ax^* = b$, nous obtenons $w^* = (x^*, \lambda^*, s^*)$ solution des conditions d'optimalité. \square

Tous les détails concernant le résultat suivant se trouvent dans la référence [1]. Etablissons le résultat de convergence locale.

Théorème 2.2 *Soit β un paramètre satisfaisant l'inégalité $\beta > 2\sqrt{n}$. Supposons que les conditions d'optimalité ont une solution unique $w^* = (x^*, \lambda^*, s^*)$ et que la suite $\{w^k\}$ engendrée par l'algorithme 2 converge vers w^* . Alors $\{\tau_k\}$ converge globalement linéairement et localement de manière quadratique vers zéro.*

L'observation centrale pour prouver la convergence du théorème 2.2 est que la suite des matrices Jacobiennes $\Theta'(w^k, \tau_k)$ converge vers une matrice non singulière sous l'hypothèse du théorème 2.2. En effet, comme mentionné en [1], la convergence de cette suite vers une matrice Jacobienne non singulière est équivalente à l'unicité de la solution des conditions d'optimalité.

2.2 Choix du point de départ et Critère d'arrêt.

Le point initial $w^0 = (x^0, \lambda^0, s^0)$ est construit comme suit :

1. Résoudre $AA^T y = b$ en utilisant la factorisation de Cholesky pour calculer $y^0 \in \mathbb{R}^m$.
2. Poser $x^0 := A^T y^0$.
3. Résoudre $AA^T \lambda = Ac$ en utilisant la factorisation de Cholesky pour calculer $\lambda^0 \in \mathbb{R}^m$.
4. Poser $s^0 := c - A^T \lambda^0$.

Notons que notre point de départ est admissible dans le sens où il vérifie les équations linéaires $A^T \lambda + s = c$ et $Ax = b$. De plus, le paramètre de régularisation initial prend la valeur

$$\tau_0 := \|\phi(x^0, s^0)\|_\infty,$$

i.e., τ_0 correspond au résidu initial des conditions d'optimalité. Pour garantir que $\phi(x^0, s^0, \tau_0) \leq 0$, nous devons augmenter la valeur de τ_0 de façon à satisfaire :

$$\tau_0 \geq \sqrt{x_i^0 s_i^0} \quad \forall i \in \{1, \dots, n\} \text{ où } x_i^0 > 0, s_i^0 > 0.$$

Regardons maintenant le critère d'arrêt.

Nous terminons l'itération lorsqu'une des conditions suivantes est satisfaite :

1. $\tau_k < 10^{-4}$,
2. $\|\Phi(w^k)\|_\infty < 10^{-4}$ ou
3. $\|\Phi(w^k)\|_\infty < 10^{-3}$ et $\|\Phi(w^k)\|_\infty / \|\Phi(w^0)\|_\infty < 10^{-6}$.

Le paramètre de centrage $\hat{\sigma}_k$ est quant à lui choisi de la façon suivante :

Posant $\hat{\sigma}_{\min} = 0.4$, $\hat{\sigma}_{\max} = 0.6$, $\gamma = 0.1$, nous commençons avec $\hat{\sigma}_0 = 0.5$ et mettons à jour $\hat{\sigma}_k$:

$$\hat{\sigma}_{k+1} := \min\{\hat{\sigma}_{\max}, \hat{\sigma}_k + \gamma\}$$

quand le pas Prédicteur est réussi, i.e., quand $\hat{w}^{k+1} = w^k + \Delta w^k$. Sinon, posons

$$\hat{\sigma}_{k+1} := \max\{\hat{\sigma}_{\min}, \hat{\sigma}_k - \gamma\}.$$

Par cette mise à jour, tous les paramètres de centrage appartiendront bien à l'intervalle $[\hat{\sigma}_{\min}, \hat{\sigma}_{\max}]$.

Les autres paramètres de l'Etape 0 sont choisis tels que

$$\rho = 0.79 \text{ et } \beta := \|\phi(x^0, s^0, \tau_0)\| / \tau_0.$$

Chapitre 3

Méthodes de type régulières améliorées pour la solution de programmes linéaires.

L'idée de base de cette méthode est de résoudre les conditions d'optimalité. Pour cela, nous exprimons les conditions d'optimalité comme un système non linéaire d'équations. Ce système étant non différentiable en général, nous l'approximerons par un système régulier d'équations. Nous appliquerons ensuite une méthode de Newton à ce système. La méthode présentée pourra être vue comme un cas particulier de la méthode du chapitre précédent.

3.1 Description de l'algorithme.

Le but de cette section est d'établir un algorithme et de l'analyser. Pour cela, soit $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction minimum déjà rencontrée, i.e.,

$$\varphi(a, b) := 2 \min\{a, b\}.$$

Définissons

$$\Phi(w) := \Phi(x, \lambda, s) := \begin{pmatrix} A^T \lambda + s - c \\ Ax - b \\ \phi(x, s) \end{pmatrix},$$

où $\phi(x, s) := (\varphi(x_1, s_1), \dots, \varphi(x_n, s_n))^T \in \mathbb{R}^n$.

Comme φ est une fonction NCP, i.e.,

$$\varphi(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0, ab = 0,$$

l'équivalence suivante est une conséquence immédiate de la définition de la fonction Φ :

$$\begin{aligned} w^* = (x^*, \lambda^*, s^*) \text{ résout les conditions d'optimalité (1.3)} \\ \text{ssi} \\ w^* \text{ est solution de } \Phi(w) = 0. \end{aligned}$$

Or, le système $\Phi(w) = 0$ est non différentiable. Par conséquent, soit $\varphi_\tau : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction régulière minimum

$$\varphi_\tau(a, b) := a + b - \sqrt{(a - b)^2 + 4\tau^2}$$

où $\tau > 0$ est un paramètre régulier. Alors, nous définissons

$$\Phi_\tau(w) := \Phi_\tau(x, \lambda, s) := \begin{pmatrix} A^T \lambda + s - c \\ Ax - b \\ \phi_\tau(x, s) \end{pmatrix}$$

où $\phi_\tau(x, s) := (\varphi_\tau(x_1, s_1), \dots, \varphi_\tau(x_n, s_n))^T \in \mathbb{R}^n$.

Par la référence [12], nous voyons que la relation suivante est satisfaite :

$$\begin{aligned} w_\tau \text{ est solution des conditions de la trajectoire centrale (1.18)} \\ \text{ssi} \\ \Phi_\tau(w_\tau) = 0, \end{aligned}$$

i.e., nous obtenons une formulation de la trajectoire centrale comme un système non-linéaire et différentiable d'équations.

Jusqu'à présent, nous avons considéré τ comme un paramètre mais il semble parfois plus utile de voir τ comme une variable indépendante. Pour éviter les ambiguïtés de notations, nous écrivons

$$\theta(x, s, \tau) := \phi_\tau(x, s).$$

Par ailleurs, nous aurons besoin de la fonction $\Theta : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times [0, \infty) \rightarrow \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}$ définie par

$$\Theta(w, \tau) := \Theta(x, \lambda, s, \tau) := \begin{pmatrix} A^T \lambda + s - c \\ Ax - b \\ \theta(x, s, \tau) \\ \tau \end{pmatrix}.$$

Remarquons que la fonction $\Theta(w, \tau)$ est différente de la fonction $\Phi_\tau(w)$ car nous avons ajouté une ligne pour obtenir la fonction Θ . Puisque l'équation $\Theta(w, \tau) = 0$ implique que $\tau = 0$, nous obtenons l'équivalence

$$w^* = (x^*, \lambda^*, s^*) \text{ résout (1.3) ssi } (w^*, 0) \text{ est solution de } \Theta(w, \tau) = 0.$$

Nous obtenons une nouvelle expression des conditions d'optimalité où τ est une variable indépendante.

Il nous semble important d'exploiter la relation entre la méthode de Newton appliquée au système $\Phi_\tau(w) = 0$ et appliquée au système $\Theta(w, \tau) = 0$. Considérons, pour commencer, le système $\Phi_\tau(w) = 0$. Nous supposons que $w^k = (x^k, \lambda^k, s^k)$ soit l'itéré et que $\tau_k > 0$ soit la valeur du paramètre de régularisation. Alors, nous définissons

$$w^{k+1} := w^k + t_k \Delta w^k$$

pour une longueur de pas $t_k > 0$ adéquate, où le vecteur correction $\Delta w^k = (\Delta x^k, \Delta \lambda^k, \Delta s^k)$ est solution du système linéaire d'équations

$$\Phi'_{\tau_k}(w^k) \Delta w^k = -\Phi_{\tau_k}(w^k).$$

Tenant compte de la définition de Φ_τ , cette équation est équivalente à

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ D_{a,\tau}^k & 0 & D_{b,\tau}^k \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{pmatrix} = \begin{pmatrix} -A^T \lambda^k - s^k + c \\ -Ax^k + b \\ -\phi_{\tau_k}(x^k, s^k) \end{pmatrix} \quad (3.1)$$

où

$$D_{a,\tau}^k := \text{diag} \left(\frac{\partial \varphi_{\tau_k}}{\partial a}(x_1^k, s_1^k), \dots, \frac{\partial \varphi_{\tau_k}}{\partial a}(x_n^k, s_n^k) \right) \in \mathbb{R}^{n \times n}$$

et

$$D_{b,\tau}^k := \text{diag} \left(\frac{\partial \varphi_{\tau_k}}{\partial b}(x_1^k, s_1^k), \dots, \frac{\partial \varphi_{\tau_k}}{\partial b}(x_n^k, s_n^k) \right) \in \mathbb{R}^{n \times n}.$$

D'autre part, si nous appliquons la méthode de Newton au système $\Theta(w, \tau) = 0$, nous devons résoudre une équation comme

$$\Theta'(w^k, \tau_k) \begin{pmatrix} \Delta w \\ \Delta \tau \end{pmatrix} = -\Theta(w^k, \tau_k)$$

à chaque itération, où nous prenons les dérivées en w et en τ . D'où, le système devient

$$\begin{pmatrix} 0 & A^T & I & 0 \\ A & 0 & 0 & 0 \\ D_{a,\tau}^k & 0 & D_{b,\tau}^k & d_\tau^k \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \\ \Delta \tau \end{pmatrix} = \begin{pmatrix} -A^T \lambda^k - s^k + c \\ -Ax^k + b \\ -\theta(x^k, s^k, \tau_k) \\ -\tau_k \end{pmatrix} \quad (3.2)$$

où

$$d_\tau^k := \left(\frac{\partial \theta}{\partial \tau}(x_1^k, s_1^k, \tau_k), \dots, \frac{\partial \theta}{\partial \tau}(x_n^k, s_n^k, \tau_k) \right)^T \in \mathbb{R}^n.$$

Poussés par des considérations semblables à celles relatives aux méthodes des points intérieurs (voir [18]), nous considérons une généralisation du système (3.2) et remplaçons le paramètre τ_k de la dernière ligne du membre de droite de (3.2) par $\sigma_k \tau_k$ pour un nombre $\sigma_k \in [0, 1]$, i.e., nous résolvons

$$\begin{pmatrix} 0 & A^T & I & 0 \\ A & 0 & 0 & 0 \\ D_{a,\tau}^k & 0 & D_{b,\tau}^k & d_\tau^k \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \\ \Delta \tau \end{pmatrix} = \begin{pmatrix} -A^T \lambda^k - s^k + c \\ -Ax^k + b \\ -\theta(x^k, s^k, \tau_k) \\ -\sigma_k \tau_k \end{pmatrix} \quad (3.3)$$

(quand $\sigma_k = 1$, nous avons (3.2)). Constatons que θ ne subit pas cette modification. Afin d'avoir des notations plus compactes pour le système linéaire (3.3), introduisons la fonction

$$\Theta_\sigma(w, \tau) := \begin{pmatrix} A^T \lambda + s - c \\ Ax - b \\ \theta(x, s, \tau) \\ \sigma \tau \end{pmatrix}$$

où l'indice σ indique la dépendance de Θ par rapport au paramètre σ . Alors, le système linéaire (3.3) peut être mis sous la forme suivante :

$$\Theta'(w^k, \tau_k) \begin{pmatrix} \Delta w \\ \Delta \tau \end{pmatrix} = -\Theta_\sigma(w^k, \tau_k).$$

Notons que le système (3.3) nous donne

$$\Delta \tau_k = -\sigma_k \tau_k. \quad (3.4)$$

En remplaçant cette expression dans les autres équations de (3.3), nous obtenons

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ D_{a,\tau}^k & 0 & D_{b,\tau}^k \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{pmatrix} = \begin{pmatrix} -A^T \lambda^k - s^k + c \\ -Ax^k + b \\ -\theta(x^k, s^k, \tau_k) + \sigma_k \tau_k d_\tau^k \end{pmatrix}. \quad (3.5)$$

D'où, nous pouvons écrire le système sous la forme

$$\Phi'_{\tau_k}(w^k) \Delta w = -\Phi_{\tau_k}(w^k) + \sigma_k \tau_k \begin{pmatrix} 0 \\ 0 \\ d_\tau^k \end{pmatrix}.$$

Le système linéaire (3.3) peut être vu comme une perturbation du système (3.1) où la perturbation n'agit que sur la troisième ligne du membre de droite.

Basé sur les notations introduites précédemment, nous énonçons notre méthode régulière Prédicteur-Correcteur.

Algorithme 3:

Etape 0 : *Initialisation.*

Choisir $w^0 := (x^0, \lambda^0, s^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ et $\tau_0 > 0$ tels que $A^T \lambda^0 + s^0 = c$, $Ax^0 = b$ et $\theta(x^0, s^0, \tau_0) \leq 0$.

Choisir $\beta \geq \|\phi_{\tau_0}(w^0)\|/\tau_0$, $\rho \in (0, 1)$, $0 < \hat{\sigma}_{\min} < \hat{\sigma}_{\max} < 1$, $\epsilon \geq 0$ et poser $k := 0$.

Etape 1 : *Critère d'arrêt.*

Si $\|\phi(w^k)\| \leq \epsilon$: Stop.

Etape 2 : *Pas Prédicteur.*

Calculer une solution $(\Delta w^k, \Delta \tau_k) = (\Delta x^k, \Delta \lambda^k, \Delta s^k, \Delta \tau_k) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}$ du système linéaire

$$\Theta'(w^k, \tau_k) \begin{pmatrix} \Delta w \\ \Delta \tau \end{pmatrix} = -\Theta(w^k, 0). \quad (3.6)$$

Si $\|\Phi(w^k + \Delta w^k)\| = 0$: STOP.

Sinon si

$$\|\theta(x^k + \Delta x^k, s^k + \Delta s^k, \tau_k)\| > \beta \tau_k$$

alors poser

$$\hat{w}^k := w^k, \hat{\tau}_k := \tau_k \text{ et } \eta_k := 1,$$

sinon calculer $\eta_k = \rho^{l_k}$ où l_k est un entier positif ou nul tel que

$$\|\theta(x^k + \Delta x^k, s^k + \Delta s^k, \rho^j \tau_k)\| \leq \beta \rho^j \tau_k \quad \forall j = 0, 1, 2, \dots, l_k \text{ et}$$

$$\|\theta(x^k + \Delta x^k, s^k + \Delta s^k, \rho^{l_k+1} \tau_k)\| > \beta \rho^{l_k+1} \tau_k,$$

et poser

$$\hat{\tau}_k := \eta_k \tau_k \text{ et } \hat{w}^k := \begin{cases} w^k & \text{si } l_k = 0 \\ w^k + \Delta w^k & \text{sinon.} \end{cases}$$

Etape 3 : Pas Correcteur.

Choisir $\hat{\sigma}_k \in [\hat{\sigma}_{\min}, \hat{\sigma}_{\max}]$ et calculer une solution $(\Delta \hat{w}^k, \Delta \hat{\tau}_k)$
 $= (\Delta \hat{x}^k, \Delta \hat{\lambda}^k, \Delta \hat{s}^k, \Delta \hat{\tau}_k) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}$ du système linéaire

$$\Theta'(\hat{w}^k, \hat{\tau}_k) \begin{pmatrix} \Delta \hat{w} \\ \Delta \hat{\tau} \end{pmatrix} = -\Theta_{\hat{\sigma}_k}(\hat{w}^k, \hat{\tau}_k). \quad (3.7)$$

Soit $\hat{t}_k = \max\{\rho^l \mid l = 0, 1, 2, \dots\}$ tel que

$$\|\theta(\hat{x}^k + \hat{t}_k \Delta \hat{x}^k, \hat{s}^k + \hat{t}_k \Delta \hat{s}^k, (1 - \hat{\sigma}_k \hat{t}_k) \hat{\tau}_k)\| \leq \beta(1 - \hat{\sigma}_k \hat{t}_k) \hat{\tau}_k \quad (3.8)$$

et poser

$$w^{k+1} := \hat{w}^k + \hat{t}_k \Delta \hat{w}^k \text{ et } \tau_{k+1} := (1 - \hat{\sigma}_k \hat{t}_k) \hat{\tau}_k.$$

Etape 4 : Mise à jour.

Poser $k = k + 1$ et aller à l'Etape 1.

Faisons quelques commentaires sur cet algorithme.

Dans l'Etape 0, nous exigeons que le point de départ $w^0 = (x^0, \lambda^0, s^0)$ soit admissible dans le sens où les équations linéaires $A^T \lambda + s = c$ et $Ax = b$ sont satisfaites. Comme les composantes de x^0 et s^0 ne doivent pas être positives (par opposition aux méthodes des points intérieurs), il est aisé de trouver un tel point de départ.

Dans le pas Prédicteur, nous calculons, pour commencer, une direction de recherche en résolvant le système linéaire (3.6). Remarquons que le membre de droite de ce système ne dépend pas de τ . En ce qui concerne le membre de gauche, nous utilisons le Jacobien standard de la fonction perturbée $\Theta(w, \tau)$. Après avoir calculé la direction dans le système (3.6), nous essayons de réduire le paramètre de régularisation τ_k autant que possible tout en s'assurant que le pas entier reste dans un certain voisinage de la trajectoire centrale (cfr le point 3 du lemme 3.2).

Le pas Prédicteur est responsable d'une convergence locale rapide tandis que le pas Correcteur assure une convergence globale.

Avant de poursuivre notre analyse, il est important de constater que cette méthode peut être vue comme un cas particulier de la méthode du chapitre précédent. En effet, si nous prenons $\psi(\tau) = \tau$ dans le chapitre 2, nous serons dans le cas de la méthode présentée ici. Donc tous les résultats établis dans le chapitre 2 restent vrais. Par ailleurs, il existe une autre façon de prouver

la convergence et nous la développerons dans la section suivante. Mais avant toute chose, remarquons que l'algorithme 3 est bien défini.

Lemme 3.1 *Les énoncés suivants sont vérifiés pour tout $k \in \mathbb{N}$:*

1. *Les systèmes linéaires (3.6) et (3.7) ont une solution unique.*
2. *Il existe un seul η_k satisfaisant les conditions de l'Etape 2.*
3. *La longueur de pas \hat{t}_k de l'Etape 3 est définie de manière unique.*

Par conséquent, l'algorithme 3 est bien défini.

Preuve :

Pour la preuve voir le lemme 2.1 du chapitre précédent .

□

Le résultat suivant sera utile dans la prochaine section.

Lemme 3.2 *Les suites $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ et $\{\tau_k\}$ engendrées par l'algorithme 3 possèdent les propriétés suivantes :*

1. *$A^T \lambda^k + s^k = c$ et $Ax^k = b$ pour tout $k \in \mathbb{N}$.*
2. *$\tau_k = \tau_0(1 - \hat{\sigma}_0 \hat{t}_0) \eta_0 \cdots (1 - \hat{\sigma}_{k-1} \hat{t}_{k-1}) \eta_{k-1}$ pour tout $k \in \mathbb{N}$.*
3. *$\|\theta(x^k, s^k, \tau_k)\| \leq \beta \tau_k$ pour tout $k \in \mathbb{N}$.*

Preuve :

Elle est semblable à celle du lemme 2.2 du chapitre précédent.

□

3.2 Propriétés de convergence.

Nous allons établir les résultats de convergence globale et locale de l'algorithme 3. Pour cela, nous supposons que le paramètre de terminaison ϵ est égal à zéro. Nous supposons aussi que l'algorithme 3 engendre un nombre infini d'itérés w^k , i.e., que nous ne nous arrêtons pas après un nombre fini d'itérations à un point w^k satisfaisant les conditions d'optimalité (1.3).

Etablissons le résultat suivant pour nous aider à montrer que la suite $\{w^k\}$ est bornée.

Lemme 3.3 *Les suites $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ et $\{\tau_k\}$ engendrées par l'algorithme 3 vérifient l'inégalité*

$$\|\min\{x^k, s^k\}\|_\infty \leq \kappa \tau_k$$

pour tout $k \in \mathbb{N}$ où $\kappa := (2 + \beta)/2$.

Preuve :

Soit θ_i la i ème composante de la fonction θ , i.e.,

$$\theta_i(a, b, \tau) := a + b - \sqrt{(a - b)^2 + 4\tau^2}.$$

Par le lemme 1.1 du premier chapitre, nous avons l'inégalité

$$|\theta_i(a, b, 0) - \theta_i(a, b, \tau)| \leq 2\tau$$

pour tous $a, b \in \mathbb{R}$ et pour tout $\tau > 0$. En utilisant la partie 3 du lemme 3.2, il suit que

$$\begin{aligned} 2|\min\{x_i^k, s_i^k\}| &= |\theta_i(x_i^k, s_i^k, 0)| \\ &\leq |\theta_i(x_i^k, s_i^k, \tau_k)| + |\theta_i(x_i^k, s_i^k, 0) - \theta_i(x_i^k, s_i^k, \tau_k)| \\ &\leq \|\theta(x^k, s^k, \tau_k)\| + |\theta_i(x_i^k, s_i^k, 0) - \theta_i(x_i^k, s_i^k, \tau_k)| \\ &\leq (\beta + 2)\tau_k \end{aligned}$$

pour tout $k \in \mathbb{N}$ et tout $i = 1, \dots, n$. D'où

$$\|\min\{x^k, s^k\}\|_\infty \leq \kappa \tau_k$$

pour tout $k \in \mathbb{N}$, où κ est la constante de notre énoncé. □

Nous voulons maintenant montrer que la suite $\{w^k\}$ engendrée par l'algorithme 3 reste bornée s'il existe un point strictement admissible pour les conditions d'optimalité (i.e., un vecteur $\hat{w} = (\hat{x}, \hat{\lambda}, \hat{s})$ satisfaisant $A^T \hat{\lambda} + \hat{s} = c$, $A\hat{x} = b$ et $\hat{x} > 0$, $\hat{s} > 0$) et si le paramètre initial $\tau_0 > 0$ est suffisamment petit.

Proposition 3.1 *Supposons qu'il existe un point strictement admissible $(\hat{x}, \hat{\lambda}, \hat{s})$ pour les conditions d'optimalité. Supposons également que le paramètre de régularisation initial $\tau_0 > 0$ vérifie*

$$\tau_0 < \frac{1}{\kappa} \min_{i=1, \dots, n} \{\hat{x}_i, \hat{s}_i\}$$

où $\kappa := (2 + \beta)/2$ est la constante du lemme 3.3. Alors, la suite $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ déterminée par l'algorithme 3 est bornée.

Preuve :

Supposons que la suite $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ n'est pas bornée. Comme la suite $\{\tau_k\}$ est décroissante, il suit du lemme 3.3 que

$$|\min\{x_i^k, s_i^k\}| \leq \|\min\{x^k, s^k\}\|_\infty \leq \kappa \tau_k \leq \kappa \tau_0 \quad (3.9)$$

pour tout $k \in \mathbb{N}$ et pour tout $i = 1, \dots, n$. D'où, il n'existe pas d'indice $i \in \{1, \dots, n\}$ tel que $x_i^k \rightarrow -\infty$ ou $s_i^k \rightarrow -\infty$ pour une sous-suite. Par conséquent, toutes les composantes des deux suites $\{x^k\}$ et $\{s^k\}$ sont bornées inférieurement.

D'autre part, la suite $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ n'est pas bornée par hypothèse. Donc, il existe au moins une composante $i \in \{1, \dots, n\}$ telle que $x_i^k \rightarrow +\infty$ ou $s_i^k \rightarrow +\infty$ pour une sous-suite. Sinon, les deux suites $\{x^k\}$ et $\{s^k\}$ seraient bornées. Cela impliquerait que la suite $\{\lambda_k\}$ le serait aussi et donc l'hypothèse ne serait pas satisfaite.

Soit $\hat{w} = (\hat{x}, \hat{\lambda}, \hat{s}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ le point strictement admissible de notre hypothèse. Alors, en particulier, nous avons

$$A^T \hat{\lambda} + \hat{s} = c \text{ et } A\hat{x} = b.$$

Puisque, nous avons aussi

$$A^T \lambda^k + s^k = c \text{ et } Ax^k = b$$

pour tout $k \in \mathbb{N}$ par le point 1 du lemme 3.2, nous obtenons

$$A^T(\hat{\lambda} - \lambda^k) + (\hat{s} - s^k) = 0 \text{ et } A(\hat{x} - x^k) = 0 \quad (3.10)$$

en soustrayant les deux équations. Prémultipliant la première par $(\hat{x} - x^k)^T$ et tenant compte de la deuxième, nous avons

$$\sum_{i=1}^n (\hat{x}_i - x_i^k)(\hat{s}_i - s_i^k) = (\hat{x} - x^k)^T (\hat{s} - s^k) = 0. \quad (3.11)$$

Nous supposons, sans perdre de généralité, qu'il existe au moins une composante i telle que $\{x_i^k\}$ est non bornée, i.e., $\{x_i^k\}_K \rightarrow +\infty$ pour un sous ensemble adéquat $K \subseteq \mathbb{N}$ (nous pouvons transposer pour s). Définissons les différents ensembles d'indices :

$$I_x := \{i \mid \{x_i^k\}_K \text{ est non bornée}\},$$

$$I_s := \{i \mid \{s_i^k\}_K \text{ est non bornée}\},$$

$$I_b := \{i \mid \{x_i^k\}_K \text{ et } \{s_i^k\}_K \text{ sont bornées}\}.$$

Remarquons que I_x est non vide tandis que I_s et I_b peuvent être vides. Utilisant les définitions de ces trois ensembles d'indices et passant à une sous-suite si nécessaire, nous obtenons de (3.9)

$$\{x_i^k\}_K \rightarrow +\infty \text{ et } s_i^k \leq \kappa\tau_0 \quad \forall k \in K \quad \forall i \in I_x \quad (3.12)$$

et

$$\{s_i^k\}_K \rightarrow +\infty \text{ et } x_i^k \leq \kappa\tau_0 \quad \forall k \in K \quad \forall i \in I_s. \quad (3.13)$$

Par ailleurs, il existe une constante $c \in \mathbb{R}$ telle que

$$\sum_{i \in I_b} (x_i^k - \widehat{x}_i)(s_i^k - \widehat{s}_i) \leq c$$

pour tout $k \in K$. Par (3.11), nous avons

$$\begin{aligned} c &\geq \sum_{i \in I_b} (x_i^k - \widehat{x}_i)(s_i^k - \widehat{s}_i) \\ &= \sum_{i \in I_x} (x_i^k - \widehat{x}_i)(\widehat{s}_i - s_i^k) + \sum_{i \in I_s} (x_i^k - \widehat{x}_i)(\widehat{s}_i - s_i^k) \end{aligned}$$

pour tout $k \in K$. Cependant, le membre de droite n'est pas borné pour une sous-suite par (3.12) et (3.13). En effet, $\widehat{s}_i - s_i^k \geq \widehat{s}_i - \kappa\tau_0 > 0$ ($i \in I_x$) et $\widehat{x}_i - x_i^k \geq \widehat{x}_i - \kappa\tau_0$ ($i \in I_s$) par le choix de $\tau_0 > 0$. Cette contradiction nous permet de conclure que la suite $\{w^k\}$ est bornée.

□

Cette proposition assure que les itérés w^k sont bornés à condition que le paramètre de régularisation initial soit suffisamment petit.

Etablissons la convergence globale de l'algorithme 3.

Théorème 3.1 *Supposons que la suite $\{w^k\} = \{(x^k, \lambda^k, s^k)\}$ déterminée par l'algorithme 3 possède au moins un point d'accumulation. Alors, la suite $\{\tau_k\}$ converge vers zéro.*

Preuve :

Elle est semblable à celle faite dans le deuxième chapitre. Voir le théorème 2.1. (Il s'agit juste d'un changement de notations).

□

Constatons que l'existence supposée d'un point d'accumulation dans ce théorème est garantie sous les conditions de la proposition 3.1. Une conséquence de ce théorème est le résultat suivant :

Corollaire 3.1 *Chaque point d'accumulation de la suite $\{w^k\}$ engendrée par l'algorithme 3 vérifie les conditions d'optimalité.*

Preuve :

Elle est analogue à celle du chapitre 2. Voir corollaire 2.1.

□

Avant d'énoncer le résultat de convergence locale, notons l'égalité entre la direction calculée dans le pas Prédicteur de notre algorithme 3 et celle obtenue dans le pas Prédicteur du premier chapitre.

Lemme 3.4 *Le vecteur $(\Delta w^k, \Delta \tau_k)$ est solution du système linéaire (3.6) de l'algorithme 3 ssi Δw^k résout le système*

$$\Phi'_{\tau_k}(w^k) \Delta w = -\Phi_0(w^k),$$

et $\Delta \tau_k = 0$.

Preuve :

Comme le paramètre de régularisation du membre de droite du système linéaire (3.6) est nul, l'assertion est une conséquence de la discussion faite sur l'équation (3.2).

□

Pour plus de détails sur le prochain théorème, consulter la référence [8].

Théorème 3.2 *Supposons que la suite $\{w^k\}$ engendrée par l'algorithme 3 converge vers une solution strictement complémentaire des conditions d'optimalité. Supposons, de plus, que le paramètre β de l'Etape 0 de l'algorithme 3 est choisi suffisamment grand tel que $\beta > 2\sqrt{n}$. Alors, le pas Prédicteur est finalement accepté et*

$$\tau_{k+1} = O(\tau^2)$$

pour tout $k \in \mathbb{N}$ suffisamment grand, i.e., le paramètre de régularisation converge localement Q -quadratiquement vers zéro.

Il est important de constater que, sous les hypothèses de ce théorème, l'ensemble des points vérifiant les conditions d'optimalité n'est pas nécessairement un singleton.

Si, à la place de la fonction minimum, nous prenons la fonction de Fischer-Burmeister, nous pouvons montrer que tous les résultats restent vrais. Par ailleurs, pour le théorème 3.2 ce n'est pas toujours satisfait. Cela reste une question ouverte.

3.3 Choix du point de départ et Critère d'arrêt.

Le point initial $w^0 = (x^0, \lambda^0, s^0)$ est construit de la manière suivante :

1. Résoudre $AA^T y = b$ en utilisant la factorisation de Cholesky pour calculer $y^0 \in \mathbb{R}^m$.
2. Poser $x^0 := A^T y^0$.
3. Résoudre $AA^T \lambda = Ac$ en utilisant la factorisation de Cholesky pour calculer $\lambda^0 \in \mathbb{R}^m$.
4. Poser $s^0 := c - A^T \lambda^0$.

C'est le même point de départ que celui de [8]. Nous obtenons le point de départ en résolvant les deux programmes simples

$$\min \frac{1}{2} \|x^2\| \quad \text{s.c. } Ax = b$$

pour x^0 et

$$\min \frac{1}{2} \|s^2\| \quad \text{s.c. } A^T \lambda + s = c$$

pour λ^0 et s^0 . La construction du point initial garantit que les deux systèmes linéaires $Ax = b$ et $A^T \lambda + s = c$ sont satisfaits en $w^0 = (x^0, \lambda^0, s^0)$. De plus, il est aisé de constater que nous avons aussi $\theta(x^0, \lambda^0, s^0) \leq 0$ à condition de prendre

$$\tau_0 \geq \sqrt{x_i^0 s_i^0} \quad \forall i \in \{1, \dots, n\} \text{ avec } x_i^0 > 0, s_i^0 > 0.$$

Ce choix est tel que le point de départ satisfait toutes les conditions requises à l'Etape 0 de l'algorithme 3.

D'autre part, nous terminons l'itération si une des conditions suivantes est vérifiée :

1. $\tau_k < 10^{-4}$ ou
2. $\|\Phi(w^k)\|_\infty < 10^{-4}$ ou
3. $\|\Phi(w^k)\|_\infty < 10^{-3}$ et $\|\Phi(w^k)\|_\infty / \|\Phi(w^0)\|_\infty < 10^{-6}$.

Le critère 1. est utilisé en [8] et est expliqué par le fait que τ^2 joue plus ou moins le rôle du saut de dualité μ dans les méthodes des points intérieurs (cfr. [2]). Or la valeur de μ pour l'arrêt est 10^{-8} . Le critère 2. désigne l'erreur absolue mesurant le résidu total $\|\Phi(w^k)\|_\infty$. Par ailleurs, le critère 3. est un 'mélange' entre une forme plus faible de cette erreur absolue et l'erreur relative au k ième résidu $\|\Phi(w^k)\|_\infty$ où $\|\Phi(w^0)\|_\infty$ est le résidu initial.

Les paramètres restants sont initialisés comme suit :

$$\rho = 0.9, \beta := \|\Phi_{\tau_0}(w^0)\| / \tau_0.$$

Pour terminer, le paramètre $\hat{\sigma}_k$ de l'Etape 3 de l'algorithme 3 vaut 0.5.

Annexe A

Quelques définitions et résultats.

Propriétés des matrices.

Soient $A, B \in \mathbb{R}^{n \times n}$ deux matrices symétriques. Alors leur somme $A + B$ est aussi symétrique. De plus, leur carré A^2 l'est également.

Soient $A, B \in \mathbb{R}^{n \times n}$ deux matrices. Si A est définie positive et B semi-définie positive, leur somme sera définie positive. Par ailleurs, la racine carrée de A est encore définie positive et celle de B est semi-définie positive.

Soient A et $B \in \mathbb{R}^{n \times n}$, ces deux matrices sont semblables ssi $A = P^{-1}BP$ où P est une matrice non singulière. De plus, A et B possèdent les mêmes valeurs propres.

Si la matrice BA est semblable à la matrice $B^{1/2}AB^{1/2}$ et que les matrices A et B sont symétriques définies positives, alors les valeurs propres de BA sont réelles et strictement positives.

En effet, il suffit de montrer que les valeurs propres de $B^{1/2}AB^{1/2}$ sont réelles et strictement positives. Par le théorème [9, thm 4.2.1], nous avons $B^{1/2}AB^{1/2}$ est définie positive. Par [10, page 469], nous savons que $B^{1/2}AB^{1/2}$ est symétrique et une matrice symétrique a ses valeurs propres réelles.

Soient E et F deux matrices symétriques définies positives qui commutent. Comme E et F commutent, $FE^{-1} = E^{-1}F$. Montrons que $E^{-1}F$ est définie positive. Puisque E^{-1} et F commutent, E^{-1} et F possèdent les mêmes vecteurs propres. Il existe donc une matrice orthogonale Q telle que

$$F = Q\Lambda Q^T = Q\Lambda^{1/2}Q^T Q\Lambda^{1/2}Q^T = F^{1/2}F^{1/2}$$

et

$$E^{-1} = Q\Gamma Q^T.$$

D'où,

$$F^{1/2} E^{-1} F^{1/2} = Q \Lambda^{1/2} \Gamma \Lambda^{1/2} Q^T$$

où $\Lambda^{1/2} \Gamma \Lambda^{1/2}$ est définie positive. Il suit que $E^{-1} F$ est aussi définie positive.

Soit $A = [a_{ij}] \in \mathbb{R}^{n \times n}$. Alors, la trace de A possède les propriétés suivantes :

1. $tr(AB) = tr(BA)$,
2. $tr(A^T A) = \|A\|^2$,
3. $tr(A) = \sum_{i=1}^n n \lambda_i$ où λ_i désigne une valeur propre de A

pour toute matrice $B \in \mathbb{R}^{n \times n}$.

Opérateur de Lyapunov.

En ce qui concerne l'opérateur de Lyapunov L_E veuillez consulter la référence [11].

Quelques propriétés des normes.

Nous sommes toujours dans le cas où A est une matrice $\in \mathbb{R}^{n \times n}$.

1. $\|AB\| \leq \|A\| \|B\|$ où la matrice $B \in \mathbb{R}^{n \times n}$,
2. $\|A^2\| \leq \|A\|^2$,
3. $\|Q^T A Q\| = \|A\|$ car la norme est invariante sous transformation orthogonale.

Calcul différentiel.

Soit $f : dom(f) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction. Soient $\langle ., . \rangle$ est le produit scalaire sur \mathbb{R}^n et $\| . \|$ la norme induite par le produit scalaire. La fonction f est dite différentiable (au sens de Fréchet) en $x \in dom(f)$ s'il existe un vecteur $g(x)$ satisfaisant

$$\lim_{\|\Delta x\| \rightarrow 0} \frac{f(x + \Delta x) - f(x) - \langle g(x), \Delta x \rangle}{\|\Delta x\|} = 0.$$

Le vecteur $g(x)$ est le gradient de f en x en ce qui concerne $\langle ., . \rangle$.

Si nous considérons la fonction ϕ de notre texte,

$$\nabla \phi(X, S, \tau)(U, V, \mu) = \lim_{t \rightarrow 0} \frac{\phi((X, S, \tau) + t(U, V, \mu)) - \phi(X, S, \tau)}{t}.$$

Explication de certaines notations.

$$g(h) = o(\|h\|) \text{ quand } h \rightarrow 0 \text{ ssi } \lim_{h \rightarrow 0} \frac{g(h)}{\|h\|} = 0.$$

$$\|x\| = O(\|v\|) \text{ ssi } \|x\| \leq c\|v\|.$$

Différence entre convergence globale et convergence locale.

Convergence globale : l'algorithme converge vers une solution à partir de n'importe quel point de départ.

Convergence locale : la vitesse à laquelle l'algorithme converge une fois proche de la solution.

Annexe B

Transformation de matrices en vecteurs.

Transformation de matrices en vecteurs.

Dans le cas d'une matrice générale $A \in \mathbb{R}^{n \times n}$, nous utilisons la fonction $vec: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n^2}$ définie par

$$vec(A) := (a_{11}, a_{21}, \dots, a_{n1}, a_{12}, a_{22}, \dots, a_{nn})^T \in \mathbb{R}^{n^2},$$

i.e., vec empile les colonnes de A dans un vecteur de longueur n^2 .

Pour une matrice A symétrique, il nous suffit de considérer la partie triangulaire inférieure de A . La transformation correspondante est possible grâce à la fonction $svec: \mathcal{S}^{n \times n} \rightarrow \mathbb{R}^{n(n+1)/2}$ donnée par

$$svec(A) := (a_{11}, \sqrt{2}a_{21}, \dots, \sqrt{2}a_{n1}, a_{22}, \sqrt{2}a_{32}, \dots, a_{nn})^T \in \mathbb{R}^{\frac{n(n+1)}{2}}.$$

La raison du facteur $\sqrt{2}$ devant les éléments non diagonaux est due au fait que c'est consistant avec le produit scalaire, i.e.,

$$A \bullet B = svec(A)^T svec(B) \forall A, B \in \mathcal{S}^{n \times n}. \quad (\text{B.1})$$

Transformation du produit matriciel.

Définissons le *produit de Kronecker* de deux matrices quelconques $G, K \in \mathbb{R}^{n \times n}$ par

$$G \otimes K := [g_{ij}K] \in \mathbb{R}^{n^2 \times n^2}.$$

De plus, il est aisé de vérifier que

$$(G \otimes K)vec(H) = vec(KHG^T) \quad (H \in \mathbb{R}^{n \times n}).$$

Dans le cas de matrices symétriques, nous avons

$$(G \otimes_S K) \text{svec}(H) := \frac{1}{2} \text{svec}(KHG^T + GHK^T) \quad (S \in \mathcal{S}^{n \times n}). \quad (\text{B.2})$$

La preuve du lemme suivant se trouve dans l'appendice du papier [16].

Lemme B.1 *Le produit de Kronecker symétrique \otimes_S défini par (B.2) possède les propriétés suivantes :*

1. $G \otimes_S K = K \otimes_S G$.
2. $G \otimes I$ est symétrique $\Leftrightarrow G$ l'est.
3. $(G \otimes_S K)(H \otimes_S L) = \frac{1}{2}(GH \otimes_S KL + GL \otimes_S KH)$.
4. Si G et K sont symétriques définies positives, alors $G \otimes_S K$ l'est aussi.

Annexe C

Tableaux comparatifs pour les trois méthodes linéaires.

Le premier tableau est celui des résultats obtenus pour certains problèmes lors de l'application de la première méthode linéaire développée. Pour plus de renseignements, voir la référence [8].

Le deuxième tableau est celui des résultats obtenus lorsque nous réalisons une mise à jour plus flexible du paramètre τ . Pour un complément d'informations voir référence [7].

Le troisième tableau est celui relatif à la dernière méthode développée. (cfr [5]).

Explication des notations utilisées pour les trois tableaux :

1. tableau 1.

- . **problem** : le nom du problème test dans la collection **netlib**
- . m : le nombre de lignes
- . n : le nombre de colonnes
- . k : le nombre d'itérations
- . J : le nombre de pas prédictors réguliers Jacobien acceptés
- . S : le nombre de pas prédictors réguliers acceptés
- . τ_f : la valeur de τ_k à l'itéré final
- . $\|\Phi(w^f)\|_\infty$: la valeur de $\|\Phi(w^k)\|_\infty$ à l'itéré final
- . **objectif primal** : la valeur de la fonction objectif primale à l'itéré final.

Nous nous arrêtons quand le nombre d'itérations dépasse 150 ou si la longueur du pas correcteur devient trop petite.

2. tableau 2.

- . `problem` : le nom du problème test dans la collection **netlib**
- . `m` : le nombre de contraintes d'égalités
- . `n` : le nombre de variables
- . `k` : le nombre d'itérations
- . `P` : le nombre de pas prédictors acceptés
- . `τ_f` : la valeur de τ_k à l'itéré final
- . `$\|\Phi(w^f)\|_\infty$` : la valeur de $\|\Phi(w^k)\|_\infty$ à l'itéré final
- . `primal objective` : la valeur de la fonction objectif primale à l'itéré final.

Les nombres entre parenthèses dans la colonne `k` correspondent au nombre d'itérations pour le troisième algorithme.

3. tableau 3.

- . `problem` : le nom du problème test dans la collection **netlib**
- . `m` : le nombre de contraintes d'égalités
- . `n` : le nombre de variables
- . `k` : le nombre d'itérations
- . `P` : le nombre de pas prédictors acceptés
- . `τ_f` : la valeur de τ_k à l'itéré final
- . `$\|\Phi(w^f)\|_\infty$` : la valeur de $\|\Phi(w^k)\|_\infty$ à l'itéré final
- . `primal objective` : la valeur de la fonction objectif primale à l'itéré final.

Numerical results for the method 1

| problem | m | n | k | J | S | τ_f | $\ \Phi(w')\ _\infty$ | primal obj |
|----------|------|-------|-----|-----|-----|------------|-----------------------|-------------------|
| degen2 | 444 | 757 | 6 | 5 | 4 | 1.2860e-05 | 2.7119e-04 | -1.4351780027e+03 |
| degen3 | 1503 | 2604 | 9 | 5 | 6 | 1.3941e-05 | 3.1761e-04 | -9.8729401032e+02 |
| e226 | 220 | 469 | 11 | 7 | 9 | 6.9197e-05 | 2.2793e-03 | -1.8751926034e+01 |
| etamacro | 357 | 692 | 16 | 4 | 15 | 4.7998e-05 | 7.1797e-04 | -7.5571524102e+02 |
| ffm300 | 501 | 1005 | 33 | 6 | 23 | 3.7494e-05 | 9.5605e-04 | 5.5567957806e+05 |
| finnis | 492 | 1014 | 21 | 3 | 20 | 5.9699e-05 | 1.3796e-03 | 1.7279109279e+05 |
| fit1d | 24 | 1049 | 27 | 6 | 12 | 1.1214e-08 | 6.9938e-07 | -9.1463780924e+03 |
| fit1p | 627 | 1677 | — | — | — | — | — | — |
| fit2d | 25 | 10524 | 13 | 6 | 6 | 8.0765e-06 | 1.5959e-03 | -6.8464293130e+04 |
| fit2p | 3000 | 13526 | — | — | — | — | — | — |
| forplan | 135 | 463 | 34 | 3 | 17 | 1.2676e-05 | 4.2004e-04 | -6.6421823582e+02 |
| ganges | 1137 | 1534 | 18 | 7 | 16 | 2.4600e-05 | 1.1170e-03 | -1.0958573613e+05 |
| gfrd-pac | 600 | 1144 | 30 | 5 | 11 | 3.0370e-05 | 7.2264e-04 | 6.9022359998e+06 |
| israel | 174 | 316 | 29 | 5 | 15 | 1.8046e-06 | 4.9876e-05 | -8.9664482215e+05 |
| kb2 | 43 | 68 | 19 | 7 | 4 | 7.8250e-05 | 3.0236e-04 | -1.7499001320e+03 |
| lotfi | 151 | 364 | 38 | 3 | 18 | 5.7195e-05 | 7.7849e-04 | -2.5264485771e+01 |
| maros | 835 | 1921 | 57 | 5 | 25 | 2.2754e-05 | 1.8723e-03 | -5.8063192480e+04 |
| modszk1 | 686 | 1622 | 31 | 9 | 16 | 6.8418e-07 | 4.7751e-05 | 3.2061972882e+02 |
| neam | 654 | 2922 | 32 | 3 | 28 | 8.0923e-05 | 4.8817e-03 | 1.4076036486e+07 |
| perold | 625 | 1530 | 58 | 3 | 24 | 5.3991e-05 | 2.4544e-03 | -9.3806515651e+03 |
| pilot | 1441 | 4657 | 43 | 3 | 23 | 4.4615e-05 | 5.5343e-03 | -5.5748462201e+02 |
| pilotja | 924 | 2044 | 98 | 1 | 27 | 9.6231e-05 | 3.1893e-03 | -6.1120672246e+03 |
| pilotwe | 722 | 2930 | 34 | 1 | 28 | 6.9184e-05 | 2.4072e-03 | -2.7201100292e+06 |
| pilot4 | 402 | 1173 | 41 | 3 | 22 | 4.7978e-05 | 1.3641e-03 | -2.5811349175e+03 |
| pilotnov | 951 | 2242 | 47 | 4 | 16 | 2.8005e-05 | 2.4620e-03 | -4.4972761883e+03 |
| recipe | 85 | 177 | 7 | 4 | 5 | 6.9831e-06 | 9.1921e-05 | -2.6681600085e+02 |
| sc105 | 105 | 163 | 9 | 4 | 7 | 2.8833e-05 | 9.1980e-05 | -5.2202055953e+01 |
| scagr7 | 129 | 185 | 15 | 6 | 8 | 5.8858e-07 | 1.1136e-05 | -2.3313898243e+06 |
| scfxm1 | 322 | 592 | 15 | 5 | 12 | 9.1602e-05 | 3.7990e-03 | 1.8416759024e+04 |
| scfxm2 | 644 | 1184 | 19 | 5 | 11 | 2.1761e-06 | 1.4029e-04 | 3.6660261585e+04 |
| scfxm3 | 966 | 1776 | 19 | 7 | 14 | 6.8249e-07 | 2.8367e-05 | 5.4901254550e+04 |
| scorpion | 375 | 453 | 15 | 3 | 10 | 2.9324e-06 | 7.2669e-05 | 1.8781248227e+03 |
| scrs8 | 485 | 1270 | 13 | 4 | 11 | 8.5648e-05 | 3.9037e-03 | 9.0429695222e+02 |

Numerical results for Algorithm 2

| problem | m | n | k | P | τ_f | $\ \Phi(w^f)\ _\infty$ | primal objective |
|----------|------|-------|---------|----|----------|------------------------|------------------|
| degen2 | 444 | 757 | 10 (23) | 10 | 1.9e-03 | 2.901e-05 | -1.43517800e+03 |
| degen3 | 1503 | 2604 | 10 (16) | 10 | 9.2e-04 | 7.742e-05 | -9.87294001e+02 |
| e226 | 198 | 429 | 14 (27) | 13 | 2.4e-04 | 1.889e-05 | -2.58649291e+01 |
| etamacro | 334 | 669 | 20 (26) | 13 | 8.8e-05 | 1.625e-03 | -7.55715232e+02 |
| fftf800 | 322 | 826 | 28 (36) | 17 | 1.2e-03 | 5.876e-04 | 5.55679564e+05 |
| finnis | 438 | 935 | 20 (31) | 17 | 2.0e-03 | 7.843e-04 | 1.72791066e+05 |
| fit1d | 24 | 1049 | 14 (20) | 14 | 2.7e-03 | 8.491e-05 | -9.14637809e+03 |
| fit1p | 627 | 1677 | 17 (19) | 14 | 7.0e-05 | 3.469e-02 | 9.14648712e+03 |
| fit2d | 25 | 10524 | 17 (22) | 17 | 1.4e-03 | 7.494e-04 | -6.84642932e+04 |
| fit2p | 3000 | 13525 | 19 (20) | 19 | 1.5e-03 | 9.397e-05 | 6.84642933e+04 |
| forplan | 121 | 447 | 26 (28) | 17 | 2.2e-03 | 4.722e-04 | -6.64218959e+02 |
| ganges | 1113 | 1510 | 20 (25) | 19 | 2.4e-03 | 1.218e-04 | -1.09585736e+05 |
| gfrd-pnc | 590 | 1134 | 17 (23) | 15 | 3.2e-02 | 4.308e-04 | 6.90223600e+06 |
| israel | 174 | 316 | 17 (27) | 15 | 1.0e-02 | 4.732e-04 | -8.06644822e+05 |
| kb2 | 43 | 68 | 15 (32) | 10 | 1.6e-03 | 1.653e-06 | -1.74990013e+03 |
| lotfi | 133 | 346 | 23 (35) | 12 | 3.2e-03 | 7.087e-04 | -2.52647043e+01 |
| maros | 655 | 1437 | 22 (37) | 14 | 2.4e-03 | 1.738e-04 | -5.80637437e+04 |
| modszkl | 665 | 1599 | 21 (26) | 17 | 7.2e-03 | 3.330e-04 | 3.20619729e+02 |
| nesm | 654 | 2922 | 46 (52) | 9 | 4.7e-04 | 4.718e-04 | 1.40760365e+07 |
| perold | 693 | 1374 | 26 (33) | 12 | 2.1e-03 | 6.564e-04 | -9.38075527e+03 |
| pilot | 1368 | 4543 | 71 (81) | 9 | 9.0e-05 | 1.600e-02 | -5.57274205e+02 |
| pilot.ja | 810 | 1804 | 30 (76) | 14 | 7.1e-04 | 9.749e-04 | -6.11313652e+03 |
| pilot.we | 701 | 2814 | 36 (61) | 10 | 2.6e-03 | 9.981e-04 | -2.72010753e+06 |
| pilot4 | 396 | 1022 | 26 (32) | 12 | 1.7e-03 | 6.888e-04 | -2.58113924e+03 |
| pilotnov | 848 | 2117 | 15 (27) | 15 | 2.3e-03 | 4.059e-04 | -4.49727619e+03 |
| recipe | 64 | 123 | 11 (14) | 10 | 1.2e-03 | 4.205e-05 | -2.66616000e+02 |
| sc105 | 104 | 162 | 18 (19) | 13 | 1.2e-03 | 2.793e-05 | -5.22020617e+01 |
| scagr7 | 127 | 183 | 15 (19) | 14 | 3.6e-03 | 4.563e-04 | -2.33138982e+06 |
| scfxm1 | 305 | 568 | 15 (20) | 13 | 8.4e-03 | 6.230e-04 | 1.84167590e+04 |
| scfxm2 | 610 | 1136 | 18 (26) | 15 | 2.7e-03 | 1.834e-04 | 3.66602616e+04 |
| scfxm3 | 915 | 1704 | 20 (26) | 15 | 5.4e-03 | 9.098e-04 | 5.49012545e+04 |
| scorpion | 340 | 412 | 19 (21) | 14 | 2.4e-04 | 1.815e-05 | 1.87812482e+03 |
| scrs8 | 421 | 1199 | 17 (21) | 14 | 1.9e-03 | 2.169e-04 | 9.04293215e+02 |

Numerical results for Algorithm 3

| problem | m | n | k | P | τ_f | $\ \tilde{w}^f\ _\infty$ | primal objective |
|-----------|------|-------|-----|----|------------|--------------------------|-------------------|
| degen2 | 444 | 757 | 23 | 23 | 2.3842e-03 | 9.9759e-05 | -1.4351779632e+03 |
| degen3 | 1503 | 2604 | 16 | 16 | 7.9692e-04 | 5.7716e-05 | -9.8729398786e+02 |
| e226 | 220 | 469 | 27 | 25 | 2.4792e-04 | 5.3902e-05 | -1.8751928739e+01 |
| clanmacro | 357 | 692 | 26 | 13 | 1.5436e-04 | 7.3792e-05 | -7.5571522983e+02 |
| fmrx00 | 501 | 1005 | 36 | 14 | 6.2879e-03 | 8.5460e-04 | 5.5567957590e+05 |
| fuols | 492 | 1014 | 31 | 20 | 1.3882e-03 | 2.9193e-04 | 1.7279127031e+05 |
| fu1d | 24 | 1049 | 20 | 18 | 5.7480e-04 | 4.0534e-05 | -9.1463780917e+03 |
| fu1p | 627 | 1677 | 19 | 19 | 1.7472e-03 | 7.3692e-06 | 9.1463780936e+03 |
| fu2d | 28 | 10524 | 22 | 20 | 6.0248e-04 | 8.2675e-05 | -6.8464293289e+04 |
| fu2p | 3000 | 13525 | 20 | 20 | 1.2942e-03 | 3.0570e-04 | 6.8464293283e+04 |
| forptsa | 135 | 463 | 29 | 17 | 4.7394e-03 | 9.3267e-04 | -6.6421820761e+02 |
| ganges | 1137 | 1534 | 25 | 20 | 3.0644e-03 | 6.4436e-04 | -1.0958573612e+05 |
| gfrd-pnc | 600 | 1144 | 23 | 16 | 1.2681e-02 | 2.4042e-04 | 6.9022360024e+06 |
| israel | 174 | 516 | 27 | 17 | 3.9926e-03 | 2.8436e-04 | -8.9664482178e+05 |
| kb2 | 43 | 68 | 32 | 10 | 3.3160e-03 | 9.8866e-05 | -1.7499000911e+03 |
| loch | 151 | 364 | 35 | 16 | 3.7591e-03 | 8.0923e-04 | -2.5263066012e+01 |
| maros | 835 | 1921 | 37 | 12 | 3.1239e-03 | 7.1513e-04 | -3.8063742927e+04 |
| modask1 | 686 | 1622 | 26 | 17 | 1.1499e-02 | 8.3608e-04 | 3.2061981568e+02 |
| neam | 654 | 2922 | 52 | 14 | 2.9135e-04 | 2.4794e-04 | 1.4076036489e+07 |
| perold | 625 | 1530 | 33 | 14 | 3.3115e-03 | 7.3875e-04 | -9.3805322461e+03 |
| pilot | 1441 | 4657 | 81 | 14 | 1.1707e-04 | 1.4059e-04 | -5.5748445014e+02 |
| pilotja | 924 | 2044 | 76 | 14 | 9.6009e-05 | 4.3618e-03 | -6.1139052461e+03 |
| pilotwa | 722 | 2930 | 61 | 13 | 5.8582e-04 | 3.7472e-04 | -2.7201075333e+06 |
| pilot4 | 402 | 1173 | 132 | 13 | 9.5377e-05 | 6.9395e-03 | -2.5810606602e+03 |
| pilotnov | 951 | 2242 | 27 | 22 | 2.5409e-03 | 3.0714e-04 | -4.4972761775e+03 |
| recipe | 85 | 177 | 14 | 14 | 1.2207e-03 | 3.5042e-05 | -2.6661598322e+02 |
| sc105 | 105 | 163 | 19 | 13 | 1.4451e-03 | 7.7940e-05 | -5.2202033312e+01 |
| scagr7 | 129 | 165 | 19 | 18 | 3.4139e-05 | 3.3656e-04 | -2.3313898243e+06 |
| scfxm1 | 322 | 592 | 20 | 19 | 3.9750e-03 | 6.4703e-04 | 1.8416759818e+04 |
| scfxm2 | 644 | 1184 | 26 | 18 | 4.3503e-03 | 8.7736e-04 | 3.6660262213e+04 |
| scfxm3 | 966 | 1776 | 26 | 21 | 4.9124e-03 | 9.6075e-04 | 5.4901255716e+04 |
| scorpion | 375 | 453 | 21 | 20 | 2.7375e-04 | 1.8825e-05 | 1.6781248227e+03 |
| scrs8 | 485 | 1270 | 21 | 19 | 6.6791e-04 | 4.5185e-05 | 9.0429695560e+02 |

Bibliographie

- [1] J.V. BURKE and S. XU. A non-interior-predictor-corrector path following algorithm for the monotone linear complementarity problem. *Math. Programming*, 87, pages 113–130. 2000.
- [2] B. CHEN and P.T. HARKER. A non-interior-point continuation method for linear complementarity problems. *SIAM J. Matrix Anal. Appl.* 14, pages 1168–1190, 1993.
- [3] X. CHEN and P. TSENG. Non-interior continuation methods for solving semidefinite complementarity problems. Technical report, Departement of Mathematics, University of Washington, Seattle, WA, 1999.
- [4] J.E. DENNIS and R.B. SCHNABEL. Numerical methods for unconstrained optimization and nonlinear equations. Prentice-Hall, Englewood Cliffs, NJ, 1983 (reprinted by SIAM, Philadelphia, PA, 1996).
- [5] S. ENGELKE and C. KANZOW. Improved smoothing-type methods for the solution of linear programs. *Numer. Math.*, 90 (2002), pp. 487–507.
- [6] S. ENGELKE and C. KANZOW. On the solution of linear programs by jacobian smoothing methods. *Annals of Operations Research* 103 (2001) pp. 49–70.
- [7] S. ENGELKE and C. KANZOW. Predictor-corrector smoothing methods for linear programs with a more flexible update of the smoothing parameter. Preprint 162, Departement of Mathematics, University of Hamburg, Hamburg, Germany, 2001.
- [8] S. ENGELKE and C. KANZOW. Predictor-corrector smoothing methods for the solution of linear programs. Preprint 153, Departement of Mathematics, University of Hamburg, Hamburg, Germany, 2000.
- [9] G. H. GOLUB and C. F. VAN LOAN. *Matrix Computations*. JOHNS HOPKINS, third edition.
- [10] R. A. HORN and C. R. JOHNSON. *Matrix Analysis*. Cambridge University Press. Cambridge, U.K., 1985.

- [11] R. A. HORN and C. R. JOHNSON. *Topics in Matrix Analysis*. Cambridge University Press. Cambridge, U.K., 1991.
- [12] C. KANZOW. Some noninterior continuation methods for linear complementarity problems. *SIAM Journal on Matrix Analysis and Applications* 17, pages 851–868, 1996.
- [13] C. KANZOW. A new approach to continuation methods for complementarity problems with uniform p-functions. *Operations Research Letters* 20, pages 85–92, 1998.
- [14] M. KOJIMA, M. SHIDA, and S. SHINDOH. A predictor-corrector interior-point algorithm for the semidefinite linear complementarity problem using the Alizadeh-Haeberly-Overton. *SIAM J. Optimization* 9, pages 444–465, 1999.
- [15] J. SUN, D. SUN, and L. QI. From strong semismoothness of the squared smoothing matrix function to semidefinite complementarity problems. Applied mathematics report amr 00/20, School of Mathematics, University of New South Wales, Sydney, Australia, 2000.
- [16] M.J. TODD, K.C. TOH, and R.H. TUTUNCU. On the Nesterov-Todd direction in semidefinite programming. *SIAM J. Optimization* 8, pages 769–796, 1998.
- [17] P. TSENG. Merit functions for semi-definite complementarity problems. *Math. Programming*, 83, pages 159–185. 1998.
- [18] S. J. WRIGHT. Primal-dual interior-point methods. *SIAM, Philadelphia, PA*, 1997.